



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

*Estimateur de la
distance minimale
pour des processus
de diffusion ergodiques*

Eric FOURNIE
Yurii KUTOYANTS

N° 1952
Juillet 1993

PROGRAMME 6

Calcul scientifique,
modélisation et
logiciels numériques

*Rapport
de recherche*

1993

ESTIMATEUR DE LA DISTANCE
MINIMALE
POUR DES PROCESSUS DE DIFFUSION
ERGODIQUES

Minimum distance estimator
for ergodic diffusion processes

Eric FOURNIÉ
INRIA Sophia-Antipolis
B.P. 93
06902 Sophia-Antipolis Cédex
France

Yurii KUTOYANTS
Université du Maine
Département de Mathématiques
B.P. 535
72017 Le Mans Cédex
France

et
U.R.A. C.N.R.S. 1321
Université Paris VI
2, place Jussieu
75251 Paris

Abstract

We study a minimum distance estimator (MDE) for ergodic diffusion processes based on the invariant measure. The criterium to minimize is $\|F_T(\cdot) - F^\theta(\cdot)\|_{L^2_\mu}$, norm of the difference between the empirical c.d.f. $F_T(\cdot)$ and the family of theoretical invariant c.d.f. $\{F^\theta(\cdot)\}$ of the model.

First, we establish its strong consistency and its asymptotic normality in the multi-dimensional case. In the case of a one-dimensional parameter, we show that on an event of probability increasing to 1 when $T \rightarrow \infty$, the minimum distance estimator admits an expansion in powers of $T^{-1/2}$. Thereafter, we study the behaviour of the MDE when the model is misspecified and we prove its convergence to some θ^* and its asymptotic normality. We conclude by giving a functional central limit theorem for the normalized empirical process when some parameters of the invariant distribution are estimated.

We give complete examples and we show the behaviour of the estimator on numerical simulations and on real data of interest rate on french monetary market.

Résumé

Nous étudions un estimateur de la distance minimale (EMD) de type Cramér-Von Mises pour des processus de diffusion ergodiques unidimensionnels. Le critère à minimiser est $\|F_T(\cdot) - F^\theta(\cdot)\|_{L^2_\mu}$, norme de la différence entre la fonction de répartition empirique $F_T(\cdot)$ et la famille $\{F^\theta(\cdot)\}$ des fonctions de répartition invariantes du modèle.

Nous établissons dans un premier temps sa consistance forte, puis sa normalité asymptotique dans le cas multidimensionnel. Dans le cas d'un paramètre unidimensionnel, nous montrons que sur un événement de probabilité croissante vers 1 lorsque $T \rightarrow \infty$, cet estimateur admet un développement en puissance de $T^{-1/2}$. Nous étudions ensuite le cas d'un modèle mal spécifié, et nous montrons la convergence de l'EMD vers une certaine valeur θ^* et sa normalité asymptotique. Nous concluons en prouvant un théorème de limite centrale fonctionnelle pour le processus empirique dans le cas où certains paramètres de la distribution invariante sont estimés par la méthode de la distance minimale.

Nous donnons des exemples complets, et nous montrons son comportement sur des simulations numériques ainsi que sur des données réelles de taux d'intérêt à un jour issues du marché monétaire français.

1 Introduction

De nombreux auteurs, depuis les travaux fondateurs de Cramér et Von Mises, ont étudié les estimateurs de la distance minimale qui constituent une sous-classe des estimateurs du minimum de contraste (cf. D. Dacunha-Castelle et M. Duflo [3]), et qui possèdent de nombreuses bonnes propriétés. Parmi les travaux les plus récents, citons ceux de P.W. Millar qui a étudié l'estimateur de la distance minimale d'un point de vue très général dans le cas des échantillons (voir [15], [16]). Il montre que pour une grande variété de distance cet estimateur est efficace et robuste. Encore plus récemment, Y.A. Kutoyants a étudié dans [9], [12], [13] cet estimateur pour différents types de métrique dans le cas des processus de diffusion faiblement bruités. Il montre sa robustesse et son caractère localement asymptotiquement minimax, et donne un développement de cet estimateur en puissance du coefficient de diffusion sur un événement de probabilité croissante vers 1 lorsque le coefficient de diffusion tend vers 0.

Dans cet article, nous étudions les performances d'un estimateur de la distance minimale d'un paramètre multidimensionnel θ pour une famille de processus de diffusion ergodiques $\{(X_t^\theta)_{t \geq 0}\}$. Le critère à minimiser est $\|F_T(\cdot) - F^\theta(\cdot)\|_{L^2_\mu}$, où μ est une mesure de probabilité, $F_T(\cdot)$ la fonction de répartition empirique et $\{F^\theta(\cdot)\}$ la famille des fonctions de répartition invariantes associée au modèle de diffusions ergodiques, c'est-à-dire un estimateur de type Cramér-Von Mises. Cette méthode d'estimation a été proposée par Y.A. Kutoyants dans [9]. Elle offre une alternative intéressante aux méthodes classiques de la statistique des diffusions.

D'un point de vue théorique, nous établissons la consistance forte de cet estimateur et sa normalité asymptotique lorsque le temps $T \rightarrow \infty$. Nous montrons que l'on peut obtenir un développement de cet estimateur en puissance de $T^{-1/2}$, sur un événement de probabilité croissante vers 1 lorsque $T \rightarrow \infty$. Les coefficients de ce développement sont des fonctions multilinéaires du processus empirique normalisé Z_T^θ et des dérivées par rapport au paramètre θ de la fonction de répartition invariante F^θ du processus. Un travail similaire a déjà été effectué pour les diffusions avec petit bruit (voir Y.A. Kutoyants [9], [10], [11] ou N. Yoshida [18]), les résultats sont alors plus intéressants en menant à la notion d'efficacité d'ordre supérieur. Nous étudions ensuite le cas d'un modèle mal spécifié, et nous montrons la convergence de notre estimateur vers la valeur θ^* minimisant la distance entre la famille des fonctions de répartition invariantes du modèle et la fonction de répartition invariante de l'observation. Sa normalité asymptotique autour de cette valeur est démontrée.

H.M. Dietz et Y.A. Kutoyants proposent dans [4] un autre estimateur de la distance minimale pour des processus de diffusion ergodiques pour lequel on obtient des formules explicites dans le cas linéaire, comme pour l'estimateur du maximum de vraisemblance, mais qui n'est pas toujours consistant.

Pour la partie étude de la consistance forte et de la normalité asymptotique de notre estimateur de la distance minimale dans le cas bien spécifié, nous avons adapté des schémas

généraux de démonstration dus à P.W. Millar [15]. Nous renvoyons à D.Talay [17] pour l'aspect simulation numérique des processus de diffusion.

2 Consistance et normalité asymptotique

Soit $\theta = (\theta^1, \dots, \theta^k) \in \Theta \subset \mathbb{R}^k$. Nous supposons dans la suite Θ ouvert, borné, et nous notons $|\cdot|$ et \langle, \rangle la norme euclidienne et le produit scalaire associé de \mathbb{R}^k .

Soient (Ω, \mathcal{F}, P) un espace probabilisé, $(\mathcal{F}_t)_{t \geq 0}$ une filtration de \mathcal{F} , et $(W_t)_{t \geq 0}$ un processus de Wiener standard sur \mathbb{R} , (\mathcal{F}_t) -adapté. Considérons la famille de diffusions sur \mathbb{R} , homogènes en temps, $\mathcal{F}_\theta = \{(X_t^\theta)_{t \geq 0}\}$, et supposons que $(X_t^\theta)_{t \geq 0}$ est l'unique solution de l'équation différentielle stochastique suivante :

$$dX_t^\theta = b_\theta(X_t) dt + \sigma_\theta(X_t) dW_t, \quad t \geq 0 \quad (2.1)$$

où $\forall \theta \in \Theta$, $b_\theta(x)$ est une fonction borélienne, bornée sur tout compact de A , et $\sigma_\theta^2(x)$ est une fonction continue, strictement positive.

On note \mathcal{L}^θ l'opérateur infinitésimal de la diffusion (X_t^θ) :

$$\mathcal{L}^\theta = b_\theta(x) \frac{\partial}{\partial x} + \frac{\sigma_\theta^2(x)}{2} \frac{\partial^2}{\partial x^2}$$

Posons $s_\theta(x) = \exp \left(- \int_0^x \frac{2b_\theta(z)}{\sigma_\theta^2(z)} dz \right)$, et $S_\theta(x) = \int_0^x s_\theta(z) dz$.

Si $\lim_{x \rightarrow \pm\infty} S_\theta(x) = \pm\infty$ la diffusion est récurrente de mesure invariante μ^θ de densité $\frac{1}{(\sigma_\theta^2(\cdot)s_\theta(\cdot))}$. Nous supposons dans la suite que pour chaque $\theta \in \Theta$, $(X_t^\theta)_{t \geq 0}$ est récurrente positive, i.e. $\int \frac{1}{\sigma_\theta^2(x)s_\theta(x)} dx < \infty$. La diffusion $(X_t^\theta)_{t \geq 0}$ est alors ergodique au sens :

$$\forall f \in L^1(\mu^\theta), \quad \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(X_s^\theta) ds = \int f d\mu^\theta, \quad P - p.s.$$

En vue de simplifier les énoncés des résultats, nous nous plaçons dans l'espace canonique. On note Q_θ^ν la loi induite par le processus $(X_t^\theta)_{t \geq 0}$ de loi initiale ν sur $(C(\mathbb{R}^+; \mathbb{R}), \mathcal{B} = \sigma(\pi_t, t \geq 0))$, où $\{\pi_t, t \geq 0\}$ sont les fonctions coordonnées, et ${}_T Q_\theta^\nu$ est sa restriction à $(C([0, T]; \mathbb{R}), \mathcal{B}_T)$, où \mathcal{B}_T est la trace de \mathcal{B} sur $C([0, T]; \mathbb{R})$. $E_{\nu, \theta}$ désigne l'espérance sous Q_θ^ν , et on note $Q_\theta^{x_0}$ et $E_{x_0, \theta}$ la loi du processus et l'espérance calculée sous cette loi, lorsque la mesure initiale est concentrée en x_0 .

Nous considérons donc l'espace statistique \mathcal{E}_0 :

$$\mathcal{E}_0 = \{C(\mathbb{R}^+; \mathbb{R}), \mathcal{B}, (Q_\theta^\nu, (\theta, \nu) \in \Theta \times \mathcal{M}_1(\mathbb{R}))\}$$

où $\mathcal{M}_1(\mathbb{R})$ désigne l'ensemble des mesures de probabilité sur \mathbb{R} .

Soit $f : \mathbb{R} \rightarrow [0, 1]$ une fonction borélienne telle que, $\forall \theta \in \Theta$, la fonction $x \rightarrow F^\theta(x) = \int \chi_{[0,x]} \circ f d\mu^\theta$ soit continue sur $[0, 1]$. Nous supposons de plus que $F^\theta(x)$ est continue en θ uniformément en x . Posons $\tilde{f}_x^\theta = \chi_{[0,x]} \circ f - \int \chi_{[0,x]} \circ f d\mu^\theta$, $F_T(x) = T^{-1} \int_0^T \chi_{[0,x]} \circ f(\pi_s) ds$, et $F^\theta(x) = \int \chi_{[0,x]} \circ f d\mu^\theta$ et définissons, pour T fixé, le processus différence empirique $(\xi_T^\theta(x), 0 \leq x \leq 1)$ et le processus empirique normalisé $(Z_T^\theta(x), 0 \leq x \leq 1)$ par :

$$\begin{aligned}\xi_T^\theta(x) &= F_T(x) - F^\theta(x), \\ Z_T^\theta(x) &= \sqrt{T} (F_T(x) - F^\theta(x)) \\ &= \frac{1}{\sqrt{T}} \int_0^T \tilde{f}_x^\theta(\pi_t) dt.\end{aligned}$$

Soit μ une mesure de probabilité sur $[0, 1]$, et considérons l'espace de Hilbert $L_\mu^2[0, 1]$. On notera $\|\cdot\|$ et (\cdot, \cdot) la norme et le produit scalaire de $L_\mu^2[0, 1]$. $\forall T > 0$, ξ_T^θ et Z_T^θ appartiennent à $L_\mu^2[0, 1]$. Nous définissons alors l'estimateur de la distance minimale, que nous notons θ_T^* , comme suit :

$$\theta_T^* \in \arg \min_{\theta \in \Theta} \|\xi_T^\theta\|$$

Nous allons montrer dans les deux paragraphes qui suivent la consistance forte et la normalité asymptotique de cet estimateur.

Posons :

$$g(\eta) = \min_{\theta_0 \in \Theta} \min_{|\theta - \theta_0| > \eta} \|F^\theta(\cdot) - F^{\theta_0}(\cdot)\|,$$

et faisons l'hypothèse suivante d'identifiabilité des lois de probabilité invariantes de la famille \mathcal{F}_θ :

(H1) $\forall \eta > 0, g(\eta) > 0$,

Posons, d'autre part, $\partial_i F^\theta(\cdot) = \frac{\partial}{\partial \theta^i} F^\theta(\cdot)$ et $\partial F^\theta(\cdot) = (\partial_1 F^\theta(\cdot), \dots, \partial_k F^\theta(\cdot))$. Nous faisons les hypothèses suivantes :

(H2) Différentiabilité : $\forall \theta \in \Theta, F^\theta \in C_\theta^2$, et $\sup_{\theta \in \Theta} \sup_x |\partial_I F^\theta(x)| \leq M_i$, pour tout multi-indice $I \in [1, \dots, k]^i, i = 1, \dots, 2$,

(H3) Non-singularité : $\exists c_0 > 0, \forall \theta \in \Theta, \forall \eta \in \mathbb{R}^k \parallel < \partial F^\theta(\cdot), \eta > \parallel \geq c_0 |\eta|$.

L'hypothèse de non-singularité (H3) entraîne :

$$\forall \eta > 0, g(\eta) \geq c_0 \eta. \quad (2.2)$$

2.1 Consistance de l'estimateur de la distance minimale

Le théorème suivant montre la consistance forte de l'EMD :

Théorème 2.1 $\forall \theta_0 \in \Theta$, sous la loi $Q_{\theta_0}^\nu$, $\theta_T^* \rightarrow \theta_0$ presque sûrement lorsque $T \rightarrow \infty$.

Preuve L'inégalité du triangle entraîne :

$$\|F^{\theta_T^*}(\cdot) - F^{\theta_0}(\cdot)\| \leq \|\xi_T^{\theta_T^*}(\cdot)\| + \|\xi_T^{\theta_0}(\cdot)\| \leq 2\|\xi_T^{\theta_0}(\cdot)\|$$

où la deuxième inégalité vient du fait que $\|\xi_T^{\theta_T^*}(\cdot)\| = \min_{\theta \in \Theta} \|\xi_T^\theta(\cdot)\| \leq \|\xi_T^{\theta_0}(\cdot)\|$.

Le théorème ergodique montre que $\forall x \in [0, 1]$, $\xi_T^{\theta_0}(x) \rightarrow 0$, $Q_{\theta_0}^{\nu, \sigma}$ - p.s. lorsque $T \rightarrow \infty$. Par conséquent, $\|\xi_T^{\theta_0}(\cdot)\| \rightarrow 0$, $Q_{\theta_0}^{\nu, \sigma}$ - p.s., lorsque $T \rightarrow \infty$.

L'hypothèse (H1) d'identifiabilité des lois invariantes nous permet de conclure. \square

Remarque : on peut montrer de la même façon, à l'aide du théorème ergodique et des arguments utilisés dans la démonstration du théorème de Glivenko-Cantelli, la consistance forte de l'estimateur en se plaçant dans l'espace de Banach $L^\infty[0, 1]$ muni de la norme $\|f(\cdot)\|_\infty = \sup_{x \in [0, 1]} |f(x)|$.

Nous montrons maintenant que l'estimateur θ_T^* a une vitesse de convergence polynomiale de tout ordre. Nous avons besoin pour cela de faire l'hypothèse suivante :

(H4) $\forall \theta \in \Theta$, $\forall x \in [0, 1]$, l'équation de Poisson $\mathcal{L}u = -\tilde{f}_x^\theta$ admet une solution $u = G_x^\theta$ vérifiant $\forall p$ entier pair, positif, $\mu^\theta(\sigma_\theta^p \dot{G}_x^\theta(\cdot)^p) < \infty$, $\mu^\theta(G_x^{\theta p}) < \infty$.

Nous verrons au paragraphe 4 un exemple de situation où cette condition est vérifiée.

Théorème 2.2 Supposons vérifiées les hypothèses (H1), (H3), (H4), alors $\forall p$ entier pair, positif, $\exists C_p > 0$, $\forall \eta > 0$, $\forall \theta_0 \in \Theta$:

$$Q_{\theta_0}^{\mu, \theta_0}(|\theta_T^* - \theta_0| > \eta) \leq \frac{C_p}{T^{p/2} g(\eta)^p} \leq C_p c_0^{-p} \eta^{-p} T^{-p/2}, \quad \forall T > 0.$$

Preuve Pour alléger les notations, nous notons \mathbf{E} l'espérance calculée sous la loi $Q_{\theta_0}^{\mu, \theta_0}$. Nous obtenons d'abord les deux inégalités suivantes :

$$\begin{aligned} \|\xi_T^{\theta_0}(\cdot)\| &\geq \min_{|\theta - \theta_0| \leq \eta} \|\xi_T^\theta(\cdot)\|, \\ \min_{|\theta - \theta_0| > \eta} \|\xi_T^\theta(\cdot)\| &\geq \min_{|\theta - \theta_0| > \eta} \|F^\theta(\cdot) - F^{\theta_0}(\cdot)\| - \|\xi_T^{\theta_0}(\cdot)\| \geq g(\eta) - \|\xi_T^{\theta_0}(\cdot)\|. \end{aligned}$$

Par conséquent,

$$\left\{ \min_{|\theta - \theta_0| \leq \eta} \|\xi_T^\theta(\cdot)\| > \min_{|\theta - \theta_0| > \eta} \|\xi_T^\theta(\cdot)\| \right\} \subset \left\{ \|\xi_T^{\theta_0}(\cdot)\| > \frac{g(\eta)}{2} \right\}.$$

c'est-à-dire,

$$\{|\theta_T^* - \theta_0| > \eta\} \subset \left\{ \|\xi_T^{\theta_0}(\cdot)\| > \frac{g(\eta)}{2} \right\}. \quad (2.3)$$

Pour tout entier pair, $p = 2m \geq 1$, l'inégalité de Markov entraîne :

$$Q_{\theta_0}^{\mu^{\theta_0}}(|\theta_T^* - \theta_0| > \eta) \leq Q_{\theta_0}^{\mu^{\theta_0}}\left(\|\xi_T^{\theta_0}(\cdot)\| > \frac{g(\eta)}{2}\right) \leq \frac{2^p}{T^{p/2}g(\eta)^p} \mathbf{E}\|Z_T^{\theta_0}(\cdot)\|^p \quad (2.4)$$

Il reste à montrer que $\forall T > 0$, $\mathbf{E}\|Z_T^{\theta_0}(\cdot)\|^p < \infty$. Nous savons que $(Z_T^{\theta_0}(x), 0 \leq x \leq 1)$ admet la représentation suivante (voir Fournié [5], th. II.2.13)

$$Z_T^{\theta_0}(x) = \frac{1}{\sqrt{T}} \int_0^T \tilde{f}_x^{\theta_0}(\pi_t) dt = \frac{1}{\sqrt{T}} \left\{ G_x^{\theta_0}(\pi_0) - G_x^{\theta_0}(\pi_T) + \int_0^T (\sigma_{\theta_0} \dot{G}_x^{\theta_0})(\pi_t) dW_t \right\} \quad (2.5)$$

L'inégalité de Jensen entraîne :

$$\begin{aligned} \mathbf{E}\|Z_T^{\theta_0}(\cdot)\|^p &= \mathbf{E} \left[\int_0^1 \left(\frac{1}{\sqrt{T}} \int_0^T \tilde{f}_x^{\theta_0}(\pi_t) dt \right)^2 d\mu(x) \right]^{p/2} \\ &\leq T^{-p/2} \int_0^1 \mathbf{E} \left[\int_0^T \tilde{f}_x^{\theta_0}(\pi_t) dt \right]^p d\mu(x) \end{aligned}$$

On utilise alors successivement :

- l'inégalité $\forall a, b > 0$, $(a + b)^n \leq 2^{n-1}(a^n + b^n)$,
- l'inégalité sur les moments de l'intégrale stochastique par rapport au mouvement Brownien (voir par exemple I. Karatzas, S.E. Shreve [7] p 163),

$$\forall m > 1, \quad \mathbf{E} \left| \int_0^T f_t dW_t \right|^{2m} \leq (m(2m-1))^m T^{m-1} \mathbf{E} \int_0^T |f_t|^{2m} dt,$$

où $(f_t)_{t \geq 0}$ une fonctionnelle mesurable, non-anticipative, adaptée à la filtration Brownienne vérifiant $\mathbf{E} \int_0^T |f_t|^{2m} dt < \infty$,

- la stationnarité de l'observation

pour mener le calcul suivant :

$$\begin{aligned} \mathbf{E} \left[\int_0^T \tilde{f}_x^{\theta_0}(\pi_t) dt \right]^p &\leq 2^{2p-1} \mathbf{E}[G_x^{\theta_0}(\pi_0)]^p + 2^{p-1} \mathbf{E} \left[\int_0^T (\sigma_{\theta_0} \dot{G}_x^{\theta_0})(\pi_s) dW_s \right]^p \leq \\ &\leq 2^{2p-1} \mathbf{E}[G_x^{\theta_0}(\pi_0)]^p + 2^{p-1} \left(\frac{p}{2}(p-1) \right)^{p/2} T^{p/2-1} \int_0^T \mathbf{E}[(\sigma_{\theta_0} \dot{G}_x^{\theta_0})(\pi_s)]^p ds \\ &= 2^{2p-1} \mu^{\theta_0}(G_x^{\theta_0 p}) + \frac{1}{2} (2p(p-1))^{p/2} T^{p/2} \mu^{\theta_0}((\sigma_{\theta_0} \dot{G}_x^{\theta_0}(\cdot))^p) \end{aligned}$$

En vertu de l'hypothèse (H4) $\mathbb{E}\|Z_T^{\theta_0}(\cdot)\|^p$ est donc majoré par une constante dépendant de p . Nous obtenons alors grâce à (2.4) l'inégalité de gauche dans l'énoncé du théorème. L'autre inégalité est obtenue en utilisant (2.2). \square

Remarque : le résultat reste vrai pour toute loi initiale ν qui vérifie $\nu \ll \mu^{\theta_0}$ et $\frac{d\nu}{d\mu^{\theta_0}}(\cdot)$ bornée.

2.2 Normalité asymptotique de l'EMD

Dans ce paragraphe, nous montrons la normalité asymptotique de l'estimateur de la distance minimale pour la norme hilbertienne $L_\mu^2[0,1]$ en utilisant des schémas de démonstration connus (voir par exemple P.W. Millard [16]), et en les adaptant à nos besoins et nous les présentons ici dans un souci de complétude de cette étude.

Nous utiliserons dans la suite le théorème suivant ([5], chapitre II).

Théorème 2.3 *Supposons que, $\forall \theta \in \Theta$, $\forall x \in [0,1]$, l'équation de Poisson $\mathcal{L}G_x^\theta = -\tilde{f}_x^\theta$ admette une solution G_x^θ vérifiant $\mu^\theta(\sigma_\theta^2 \dot{G}_x^{\theta 2}) < \infty$. Alors quelle que soit la loi initiale ν , $\forall \theta \in \Theta$, sous Q_θ^ν , $Z_T^\theta \Rightarrow Z^\theta$ dans $D[0,1]$ où Z^θ est un processus gaussien de moyenne nulle et de fonction de covariance $c^\theta(x,y)$ donnée par :*

$$\forall x, y \in [0,1], c^\theta(x,y) = \int 4\sigma_\theta^2(u)s_\theta^2(u) \int_{-\infty}^u \frac{\tilde{f}_x^\theta(v)}{(\sigma_\theta^2 s_\theta)(v)} dv \int_{-\infty}^u \frac{\tilde{f}_y^\theta(v)}{(\sigma_\theta^2 s_\theta)(v)} dv \mu^\theta(du)$$

De plus, $Q_\theta^\nu(Z^\theta \in C[0,1]) = 1$.

On montre alors le

Théorème 2.4 *Supposons vérifiées les hypothèses (H1)-(H3) et les hypothèses du théorème 2.3, alors $\forall \theta_0 \in \Theta$, on a sous la loi $Q_{\theta_0}^\nu$:*

$$\sqrt{T}(\theta_T^* - \theta_0) \Rightarrow \mathcal{N}(0, \Gamma_{\theta_0}), \quad T \rightarrow \infty$$

où $\Gamma_{\theta_0} = J_{\theta_0}^{-1} K_{\theta_0} {}^t J_{\theta_0}^{-1}$, où t désigne la transposition, $K_{\theta_0} = (K_{\theta_0}^{ij})$, où

$$K_{\theta_0}^{ij} = \int_0^1 \int_0^1 c^{\theta_0}(x,y) \partial_i F^{\theta_0}(y) \partial_j F^{\theta_0}(x) d\mu(y) d\mu(x),$$

et $J_{\theta_0} = (\partial F^{\theta_0}(\cdot), {}^t \partial F^{\theta_0}(\cdot))$.

Preuve On écrira dans la suite $o(\cdot)$ à la place de $o_{Q_{\theta_0}^\nu}(\cdot)$. On rappelle que la notation $u = o_{Q_{\theta_0}^\nu}(v)$ signifie que pour tout $\varepsilon > 0$, $\lim_{v \rightarrow 0} Q_{\theta_0}^\nu \left(\left| \frac{u}{v} \right| > \varepsilon \right) = 0$.

Posons $r_t = \frac{6}{c_0} \|Z_t^{\theta_0}(\cdot)\|$, et définissons l'ensemble aléatoire

$$A_t = \left\{ \theta \in \Theta : |\theta - \theta_0| \leq \frac{r_t}{\sqrt{t}} = \frac{6}{c_0} \|\xi_t^{\theta_0}(\cdot)\| \right\}.$$

Définissons d'autre part un ensemble $B_{\theta_0} = \{\theta \in \Theta : |\theta - \theta_0| \leq \frac{c_0}{kM_2}\}$. Les hypothèses (H2)-(H3) entraînent pour tout $\theta \in B_{\theta_0}$:

$$\begin{aligned} \|\xi_T^\theta(\cdot)\| &\geq \| \langle \theta - \theta_0, \partial F^{\theta_0}(\cdot) \rangle \| - \|\xi_T^{\theta_0}(\cdot)\| - \frac{kM_2}{2} |\theta - \theta_0|^2 \\ &\geq \frac{1}{2} c_0 |\theta - \theta_0| - \|\xi_T^{\theta_0}(\cdot)\| \end{aligned}$$

En vertu de la définition de A_T , $\forall \varepsilon > 0$, $\exists T_0(\varepsilon) > 0$, tel que $\forall T > T_0$, $Q_{\theta_0}^\nu(A_T \subset B_{\theta_0}) > 1 - \varepsilon$. On peut alors continuer le calcul précédent et en déduire que $\forall T > T_0$, avec une probabilité plus grande que $1 - \varepsilon$:

$$\min_{\theta \in A_T^c} \|\xi_T^\theta(\cdot)\| \geq 2\|\xi_T^{\theta_0}(\cdot)\| \geq 2 \min_{\theta \in A_T} \|\xi_T^\theta(\cdot)\|$$

où A_T^c désigne le complémentaire de A_T .

Par conséquent, $\forall T > T_0$, $\theta_T^* \in A_T$ avec une probabilité plus grande que $1 - \varepsilon$:

$$\|\xi_T^{\theta_T^*}(\cdot)\| = \min_{\theta \in \Theta} \|\xi_T^\theta(\cdot)\| = \min_{\theta \in A_T} \|\xi_T^\theta(\cdot)\|$$

$$\text{i.e. } \|\theta_T^* - \theta_0\| \leq \frac{r_T}{\sqrt{T}}.$$

Il s'ensuit que, $\forall \theta \in A_T$, $\xi_T^\theta(\cdot) = \xi_T^{\theta_0}(\cdot) + \langle \theta - \theta_0, -\partial F^{\theta_0}(\cdot) \rangle + o(T^{-1/2})$.

Notons H_{θ_0} le sous-espace de $L_\mu^2[0, 1]$ engendré par la famille de fonctions $\{\partial_i F^{\theta_0}(\cdot), i = 1 \dots k\}$, et $P_{H_{\theta_0}}$ la projection sur H_{θ_0} . Comme $\forall X \in L_\mu^2[0, 1]$, $X = P_{H_{\theta_0}} X + (I - P_{H_{\theta_0}}) X$, où I désigne l'identité, en utilisant le théorème de Pythagore, on peut mener le calcul suivant :

$$\begin{aligned} \|\xi_T^{\theta_T^*}(\cdot)\|^2 &= \min_{\theta \in A_T} \|\xi_T^\theta(\cdot)\|^2 \\ &= \min_{\theta \in A_T} \|\xi_T^{\theta_0}(\cdot) + \langle \theta - \theta_0, -\partial F^{\theta_0}(\cdot) \rangle\|^2 + o(T^{-1}) \\ &= \min_{\theta \in \Theta} \|P_{H_{\theta_0}} \xi_T^{\theta_0}(\cdot) + \langle \theta - \theta_0, -\partial F^{\theta_0}(\cdot) \rangle\|^2 \\ &\quad + \|(I - P_{H_{\theta_0}}) \xi_T^{\theta_0}(\cdot)\|^2 + o(T^{-1}) \\ &= \|(I - P_{H_{\theta_0}}) \xi_T^{\theta_0}(\cdot)\|^2 + o(T^{-1}) \end{aligned}$$

De la même façon, on a

$$\begin{aligned}
& \|(I - P_{H_{\theta_0}})\xi_T^{\theta_0}(\cdot) - \xi_T^{\theta_T^*}(\cdot)\|^2 \\
&= \|(I - P_{H_{\theta_0}})\xi_T^{\theta_0}(\cdot) - (\xi_T^{\theta_0}(\cdot) + \langle \theta_T^* - \theta_0, -\partial F^{\theta_0}(\cdot) \rangle)\|^2 + o(T^{-1}) \\
&= \|\xi_T^{\theta_0}(\cdot) + \langle \theta_T^* - \theta_0, -\partial F^{\theta_0}(\cdot) \rangle\|^2 - \|(I - P_{H_{\theta_0}})\xi_T^{\theta_0}(\cdot)\|^2 + o(T^{-1}) \\
&= \|\xi_T^{\theta_T^*}(\cdot)\|^2 - \|(I - P_{H_{\theta_0}})\xi_T^{\theta_0}(\cdot)\|^2 + o(T^{-1}) \\
&= o(T^{-1})
\end{aligned}$$

On en déduit :

$$\begin{aligned}
(I - P_{H_{\theta_0}})\xi_T^{\theta_0}(\cdot) &= \xi_T^{\theta_T^*}(\cdot) + o(T^{-1/2}) \\
&= \xi_T^{\theta_0}(\cdot) + \langle \theta_T^* - \theta_0, -\partial F^{\theta_0}(\cdot) \rangle + o(T^{-1/2})
\end{aligned}$$

et par conséquent,

$$P_{H_{\theta_0}}\xi_T^{\theta_0}(\cdot) = \langle \theta_T^* - \theta_0, \partial F^{\theta_0}(\cdot) \rangle + o(T^{-1/2})$$

Il s'ensuit que :

$$\begin{aligned}
\sqrt{T}(\theta_T^* - \theta_0) &= J_{\theta_0}^{-1} \left(\partial F^{\theta_0}(\cdot), P_{H_{\theta_0}} Z_T^{\theta_0}(\cdot) \right) + o(1) \\
&= J_{\theta_0}^{-1} \left\{ \left(\partial F^{\theta_0}(\cdot), P_{H_{\theta_0}} Z_T^{\theta_0}(\cdot) \right) + \left(\partial F^{\theta_0}(\cdot), P_{H_{\theta_0}^\perp} Z_T^{\theta_0}(\cdot) \right) \right\} + o(1) \\
&= J_{\theta_0}^{-1} \left(\partial F^{\theta_0}(\cdot), Z_T^{\theta_0}(\cdot) \right) + o(1).
\end{aligned}$$

Le théorème 2.3, entraîne alors la convergence en loi, sous $Q_{\theta_0}^{\nu, \sigma}$, lorsque $T \rightarrow \infty$, de $\sqrt{T}(\theta_T^* - \theta_0)$ vers une variable gaussienne centrée. Un calcul direct donne l'expression de la matrice de covariance de cette variable aléatoire. \square

3 Développement de l'estimateur de la distance minimale

On donne ici un développement complet de l'estimateur de la distance minimale. C'est un résultat classique de la statistique, mais l'idée de développer en puissance de $\xi_T^{\theta_0}(\cdot) = F_T(\cdot) - F^{\theta_0}(\cdot)$ dans l'espace de Hilbert $L_\mu^2[0, 1]$ semble nouvelle.

Nous supposons dans ce paragraphe le paramètre θ à valeurs dans un ouvert borné de \mathbb{R} , i.e. $\theta \in \Theta = (a, b) \subset \mathbb{R}$, et nous faisons l'hypothèse suivante :

(H5) $\forall \theta \in \Theta \subset \mathbb{R}$, $F^\theta \in C_\theta^{n+2}$, et toutes les dérivées d'ordre i , $\partial^i F^\theta(\cdot)$ sont bornées uniformément en θ , $\sup_{\theta \in \Theta} \sup_x |\partial^i F^\theta(x)| \leq M_i$, pour $i = 1, \dots, n+2$.

Considérons $\min_{\theta \in \Theta} \|\xi_T^\theta\|^2$. Nous venons de voir que ce minimum est atteint aux points $\theta_T^* \in \Theta$ qui vérifient la relation $(\xi_T^{\theta_T^*}(\cdot), \partial F^{\theta_T^*}(\cdot)) = 0$. En posant $\varepsilon(\cdot) = \xi_T^{\theta_0}(\cdot)$ et $u = \theta_T^* - \theta_0$, on réécrit cette relation de la façon suivante :

$$\begin{aligned} (\xi_T^{\theta_T^*}(\cdot), \partial F^{\theta_T^*}(\cdot)) &= (\xi_T^{\theta_0}(\cdot) - (F^{\theta_0}(\cdot) - F^{\theta_T^*}(\cdot)), \partial F^{\theta_T^*}(\cdot)) \\ &= (\varepsilon(\cdot) - (F^{\theta_0+u}(\cdot) - F^{\theta_0}(\cdot)), \partial F^{\theta_0+u}(\cdot)) \\ &:= H(\varepsilon(\cdot), u) \end{aligned}$$

La fonction H ainsi définie est une fonction à valeurs réelles, définie sur $L_\mu^2[0, 1] \times \mathbb{R}$.

Pour obtenir le développement de l'estimateur θ_T^* , il suffit d'obtenir, si elle existe, la fonction $u(\varepsilon(\cdot))$, unique solution de l'équation implicite :

$$H(\varepsilon(\cdot), u(\varepsilon(\cdot))) = 0 \quad (3.6)$$

et de la développer par la formule de Taylor si elle est suffisamment régulière. Les hypothèses (H1)-(H5) nous permettront de le faire.

Nous énonçons maintenant le résultat principal de ce paragraphe.

Théorème 3.5 *Supposons vérifiées les hypothèses (H1)-(H5), alors il existe une famille d'événements $\{A(T), T \geq T_0 > 0\}$ et des variables aléatoires réelles $\psi_j^T, \psi_j, j = 1, \dots, n$, ζ_T^1, ζ_T^2 , telles que l'estimateur de la distance minimale θ_T^* admet le développement suivant :*

$$\theta_T^* = \theta_0 + \left\{ \sum_{j=1}^n T^{-j/2} \psi_j^T + T^{(-(n+1)/2+\delta)} \zeta_T^1 \right\} \chi_{\{A(T)\}} + \zeta_T^2 \chi_{\{A^c(T)\}}$$

où $\delta \in (0, \frac{1}{2})$, $|\zeta_T^1| \leq 1$, $\psi_j^T \Rightarrow \psi_j, j = 1, \dots, n$, lorsque $T \rightarrow \infty$ et

$$\begin{aligned} Q_{\theta_0}^{\mu^{\theta_0}, \sigma}(|\zeta_T^2| > T^{-\delta}) &\leq C_1(p, T_0) T^{-p(1/2-\delta)} \\ Q_{\theta_0}^{\mu^{\theta_0}, \sigma}(A^c(T)) &\leq C_2(p, T_0) T^{-p \min(\frac{1}{2}-\delta, \frac{\delta}{n+1})} \end{aligned}$$

où $C_1(p, T_0), C_2(p, T_0) > 0, p$ entier positif. On a en particulier :

$$\begin{aligned} \psi_1^T &= J^{-1}(\theta_0) \left(Z_T^{\theta_0}(\cdot), \partial F^{\theta_0}(\cdot) \right), \\ \psi_2^T &= J^{-3}(\theta_0) \left(3 \left(Z_T^{\theta_0}(\cdot), \partial F^{\theta_0}(\cdot) \right)^2 \left(\partial F^{\theta_0}(\cdot), \partial^2 F^{\theta_0}(\cdot) \right) - \right. \\ &\quad \left. - 2J(\theta_0) \left(Z_T^{\theta_0}(\cdot), \partial F^{\theta_0}(\cdot) \right) \left(Z_T^{\theta_0}(\cdot), \partial^2 F^{\theta_0}(\cdot) \right) \right) \end{aligned}$$

où l'on a posé $J(\theta_0) = (\partial F^{\theta_0}(\cdot), \partial F^{\theta_0}(\cdot))$.

Nous aurons besoin dans la suite de connaître les différentielles de tout ordre, si elles existent, de la fonction u solution de (3.6). La règle de différentiation composée nous permet d'obtenir l'expression des différentielles successives de u en fonction de celles de

H . La différentielle d'ordre j , $j \geq 1$ de la fonction u , que l'on note $u^{(j)}$, est donnée par la formule suivante :

$$u^{(j)} = \partial_u H^{1-2j} \sum_I C_I \prod_{s=1}^j \prod_{s=0}^r \partial_{\varepsilon^s u^{r-s}}^r H^{i_{r,s}} \quad (3.7)$$

où la somme est effectuée sur tous les ensembles $I = \{i_{1,0}, i_{1,1}, \dots, i_{j,j}\}$ d'entiers non négatifs satisfaisant $\sum_{r=1}^j \sum_{s=0}^r r i_{r,s} = 3j - 2$. Les C_I sont des constantes indépendantes de la fonction H .

A titre d'exemple, les deux premières différentielles de u sont :

$$\begin{aligned} u^{(1)} &= -\frac{\partial_\varepsilon H}{\partial_u H}, \\ u^{(2)} &= -\frac{\partial_{uu}^2 H \partial_\varepsilon H \partial_\varepsilon H - 2 \partial_u H \partial_{u\varepsilon}^2 H \partial_\varepsilon H + \partial_u H^2 \partial_{\varepsilon\varepsilon}^2 H}{\partial_u H^3}. \end{aligned}$$

Un calcul direct nous permet d'obtenir la forme générale des différentielles $\partial_{\varepsilon^s u^{r-s}}^r H$ de la fonction H .

$$\begin{aligned} \partial_{\varepsilon^s u^{r-s}}^r H(\varepsilon(\cdot), u) \cdot \delta_1(\cdot) \dots \delta_s(\cdot) &= 0, \quad r \geq s \geq 2, \quad \forall \delta_1(\cdot), \dots, \delta_s(\cdot) \in L_\mu^2[0, 1] \\ \partial_{\varepsilon u^{r-1}}^r H(\varepsilon(\cdot), u) \cdot \delta(\cdot) &= (\delta(\cdot), \partial^r F^{\theta_0+u}(\cdot)), \quad \forall \delta \in L_\mu^2[0, 1] \\ \partial_{u^r}^r H(\varepsilon(\cdot), u) &= \sum_{i=1}^r K_{ir} (\partial^i F^{\theta_0+u}(\cdot), \partial^{r+1-i} F^{\theta_0+u}(\cdot)) \\ &\quad + (\varepsilon(\cdot) - (F^{\theta_0}(\cdot) - F^{\theta_0+u}(\cdot)), \partial^{r+1} F^{\theta_0+u}(\cdot)) \end{aligned}$$

où les K_{ij} , $i, j = 1 \dots (n+1)$ sont des constantes indépendantes de $F^\theta(\cdot)$.

Preuve (du théorème 3.5) Soit $\delta \in (0, \frac{1}{2})$, et considérons la famille d'événements suivants :

$$\begin{aligned} A_1(T) &= \left\{ \omega : \|\xi_T^{\theta_0}(\cdot)\| < \frac{c_0^2}{M_1} T^{-\delta} - \frac{3M_2}{2} T^{-2\delta} \right\}, \\ A_2(T) &= \left\{ \omega : \|\xi_T^{\theta_0}(\cdot)\| \leq \frac{c_0^2}{2M_2} - M_1 T^{-\delta} \right\}, \\ A_3(T) &= \left\{ \omega : \|\xi_T^{\theta_0}(\cdot)\| < C_{n+1}^{-\frac{1}{n+1}} T^{-\frac{1}{2} + \frac{\delta}{n+1}} \right\}, \\ A(T) &= \cap_{i=1..3} A_i(T). \end{aligned}$$

où C_{n+1} est une constante que nous préciserons dans la suite et les constantes M_i sont définies dans l'hypothèse (H5).

Nous allons montrer que sur l'événement $A_1(T) \cap A_2(T)$, il y a existence et unicité d'une solution à l'équation implicite (3.6). Pour cela, il suffit de vérifier les conditions suivantes :

- $H(\xi_T^{\theta_0}(\cdot), -T^{-\delta}) > 0$, $H(\xi_T^{\theta_0}(\cdot), T^{-\delta}) < 0$, sur $A_1(T)$
- $\partial_u H(\xi_T^{\theta_0}(\cdot), u) < 0$ sur $A_2(T)$, $\forall u : |u| < T^{-\delta}$

La première condition assure l'existence d'une solution l'équation (3.6) et la deuxième son unicité. Les deux lemmes suivants prouvent que ces conditions sont satisfaites.

Lemme 3.6 Sur $A_1(T)$, on a $H(\xi_T^{\theta_0}(\cdot), -T^{-\delta}) > 0$ et $H(\xi_T^{\theta_0}(\cdot), T^{-\delta}) < 0$.

Preuve En effet, en revenant à la définition de H , et en utilisant la formule de Taylor, l'inégalité de Cauchy-Schwartz-Bouniakowski et les hypothèses (H2)-(H3), on obtient l'inégalité suivante pour tout $\varepsilon(\cdot) \in L_\mu^2[0, 1]$:

$$\begin{aligned} H(\varepsilon(\cdot), -T^{-\delta}) &\geq -\|\varepsilon(\cdot)\| \sup_{\theta} \|\partial F^{\theta}(\cdot)\| + T^{-\delta} \min_{\theta} \|\partial F^{\theta}(\cdot)\|^2 \\ &\quad - T^{-2\delta} \frac{3}{2} \sup_{\theta} \|\partial F^{\theta}(\cdot)\| \sup_{\theta} \|\partial^2 F^{\theta}(\cdot)\| \\ &\geq -M_1 \|\varepsilon(\cdot)\| + c_0^2 T^{-\delta} - \frac{3M_1 M_2}{2} T^{-2\delta} \end{aligned}$$

qui entraîne $H(\xi_T^{\theta_0}(\cdot), -T^{-\delta}) > 0$ sur $A_1(T)$.

De la même façon, on obtient $H(\xi_T^{\theta_0}(\cdot), T^{-\delta}) < 0$ sur $A_1(T)$. Par conséquent, il existe $u, |u| < T^{-\delta}$ tel que $H(\xi_T^{\theta_0}(\cdot), u) = 0$. \square

Lemme 3.7 Sur $A_2(T)$, $\forall u, |u| < T^{-\delta}$, $\partial_u H(\xi_T^{\theta_0}(\cdot), u) \leq -c_0^2/2$.

Preuve Pour tout $\varepsilon(\cdot) \in L_\mu^2[0, 1]$ et $u \in \mathbb{R}$, on a :

$$\partial_u H(\varepsilon(\cdot), u) = -\left(\partial F^{\theta_0+u}(\cdot), \partial F^{\theta_0+u}(\cdot)\right) + \left(\varepsilon(\cdot) - (F^{\theta_0+u}(\cdot) - F^{\theta_0}(\cdot)), \partial^2 F^{\theta_0+u}(\cdot)\right)$$

De la même façon que précédemment, on utilise les accroissements finis, l'inégalité de Cauchy-Schwartz et les hypothèses (H2)-(H3) pour mener le calcul suivant :

$$\begin{aligned} \partial_u H(\varepsilon(\cdot), u) &\leq -\min_{\theta} \|\partial F^{\theta}(\cdot)\|^2 + \|\varepsilon(\cdot)\| \sup_{\theta} \|\partial^2 F^{\theta}(\cdot)\| + \\ &\quad + |u| \sup_{\theta} \|\partial F^{\theta}(\cdot)\| \sup_{\theta} \|\partial^2 F^{\theta}(\cdot)\| \\ &\leq -c_0^2 + M_2 \|\varepsilon(\cdot)\| + M_1 M_2 |u| \end{aligned}$$

où M_1, M_2 sont les constantes données dans l'hypothèse (H5).

Donc sur $A_2(T)$, on a $\partial_u H(\xi_T^{\theta_0}(\cdot), u) \leq -c_0^2/2$, pour $u, |u| < T^{-\delta}$. \square

Nous avons donc montré l'existence et l'unicité d'une solution u à l'équation implicite (3.6) sur l'évènement $A_1(T) \cap A_2(T)$. La forme générale des différentielles $u^{(i)}$ de cette

fonction, données précédemment, et les hypothèses (H3), (H5), entraînent que les $u^{(i)}$ sont définies, continues jusqu'à l'ordre $n+1$, et bornées sur les boules $B(0, R)$, $\forall R > 0$ de $L_\mu^2[0, 1]$. Nous pouvons donc effectuer sur $A_1(T) \cap A_2(T)$ le développement de Taylor suivant :

$$\begin{aligned} u(\xi_T^{\theta_0}(\cdot)) &= u(0) + u^{(1)}(0) \cdot \xi_T^{\theta_0}(\cdot) + \frac{1}{2} u^{(2)}(0) \cdot \xi_T^{\theta_0}(\cdot) \cdot \xi_T^{\theta_0}(\cdot) + \dots + \\ &\quad + \frac{1}{n!} u^{(n)}(0) \cdot \xi_T^{\theta_0}(\cdot) \dots \xi_T^{\theta_0}(\cdot) + R_{n+1}(\xi_T^{\theta_0}(\cdot)) \end{aligned} \quad (3.8)$$

où $R_{n+1}(\xi_T^{\theta_0})$ est le reste de Taylor.

Nous pouvons calculer successivement les $u^{(j)}(0) \cdot \xi_T^{\theta_0}(\cdot) \dots \xi_T^{\theta_0}(\cdot)$, pour $j = 0 \dots n$. En particulier, pour les trois premiers termes, en posant $J(\theta_0) = (\partial F^{\theta_0}(\cdot), \partial F^{\theta_0}(\cdot))$, on obtient :

$$\begin{aligned} u(0) &= 0, \\ u^{(1)}(0) \cdot \xi_T^{\theta_0}(\cdot) &= T^{-1/2} J^{-1}(\theta_0) (Z_T^{\theta_0}(\cdot), \partial F^{\theta_0}(\cdot)) \\ &= T^{-1/2} \psi_1^T, \\ u^{(2)}(0) \cdot \xi_T^{\theta_0}(\cdot) \cdot \xi_T^{\theta_0}(\cdot) &= T^{-1} J^{-3}(\theta_0) \left(3 (Z_T^{\theta_0}(\cdot), \partial F^{\theta_0}(\cdot))^2 (\partial F^{\theta_0}(\cdot), \partial^2 F^{\theta_0}(\cdot)) \right. \\ &\quad \left. - 2J(\theta_0) (Z_T^{\theta_0}(\cdot), \partial F^{\theta_0}(\cdot)) (Z_T^{\theta_0}(\cdot), \partial^2 F^{\theta_0}(\cdot)) \right) \\ &= T^{-1} \psi_2^T. \end{aligned}$$

Les variables ψ_j^T sont des fonctions continues multi-linéaires de $Z_T^{\theta_0}(\cdot)$ que nous notons F_j . Le théorème 2.3 entraîne alors la convergence en loi des $\psi_j^T \Rightarrow \psi_j$ où $\psi_j = F_j(Z^{\theta_0})$.

Nous allons maintenant montrer que sur l'événement $A(T)$ le reste de Taylor R_{n+1} est petit devant les autres termes du développement.

Lemme 3.8 Sur $A(T)$, $|R_{n+1}(\xi_T^{\theta_0}(\cdot))| \leq T^{-(n+1)/2+\delta}$.

Preuve Comme $\|u^{(n+1)}(\varepsilon(\cdot))\|$ est bornée sur les boules $B(0, R)$, $\forall R > 0$ de $L_\mu^2[0, 1]$, en vertu de l'hypothèse (H5) par une certaine constante C_{n+1} fonction des constantes $c_0, M_i, K_{ij}, i, j = 1 \dots (n+1)$, on a

$$|R_n(\varepsilon(\cdot))| \leq C_{n+1} \|\varepsilon(\cdot)\|^{n+1}.$$

Sur $A(T)$, on a $|R_{n+1}(\xi_T^{\theta_0}(\cdot))| \leq C_{n+1} \|\xi_T^{\theta_0}(\cdot)\|^{(n+1)} \leq T^{-(n+1)/2+\delta}$. \square

Pour terminer la preuve du théorème 3.5, il reste à montrer que la probabilité de l'événement $A(T)$ tend vers 1 lorsque $T \rightarrow \infty$.

Lemme 3.9 $\forall T > T_0$, on a la majoration suivante :

$$Q_{\theta_0}^{\mu^{\theta_0}, \sigma}(A^c(T)) \leq C_2(p, T_0) T^{-p \min(\frac{1}{2}-\delta, \frac{\delta}{n+1})}$$

pour certaines constantes $C_1(p, T_0), C_2(p, T_0) > 0$, p entier pair, positif.

Preuve La preuve est alors une application immédiate de l'inégalité déduite dans le théorème 2.2 :

$$Q_{\theta_0}^{\mu^{\theta_0}, \sigma} \left(\|\xi_T^{\theta_0}(\cdot)\| > \frac{g(\eta)}{2} \right) \leq \frac{C_p}{T^{p/2} g(\eta)^p} \quad (3.9)$$

et du fait que

$$Q_{\theta_0}^{\mu^{\theta_0}, \sigma} (A^c(T)) \leq \sum_{i=1..3} Q_{\theta_0}^{\mu^{\theta_0}, \sigma} (A_i^c(T)).$$

□

On rappelle enfin (cf. théorème 2.2) que

$$Q_{\theta_0}^{\mu^{\theta_0}, \sigma} (|\theta_T^* - \theta_0| > \eta) \leq C_p c_0^{-p} \eta^{-p} T^{-p/2}$$

qui entraîne $Q_{\theta_0}^{\mu^{\theta_0}, \sigma} (|\zeta_T^2| > T^{-\delta}) \leq C_1(p, T_0) T^{-p(1/2-\delta)}$.

Ceci termine la preuve du théorème 3.5.

On déduit en particulier de ce théorème la normalité asymptotique de l'estimateur de la distance minimale avec une loi limite que nous avons déjà obtenue au paragraphe précédent pour le cas plus général d'un paramètre multidimensionnel (voir le théorème 2.4).

4 Essais numériques

4.1 Exemple 1 : estimation d'un paramètre de rappel

Nous considérons la famille $\mathcal{F}_\theta = \{(X_t^\theta)_{t \geq 0}, \theta \in \Theta = (a, b), a, b > 0\}$ de processus d'Ornstein-Uhlenbeck vérifiant les équations différentielles stochastiques

$$dX_t^\theta = -\theta X_t^\theta dt + dW_t$$

$\forall \theta > 0$, (X_t^θ) est récurrent positif, de mesure invariante $d\mu^\theta(x) = \sqrt{\frac{\theta}{\pi}} \exp(-\theta x^2) dx$.

Nous choisissons pour $f : \mathbb{R} \rightarrow [0, 1]$, la fonction $f(x) = \exp(-x^2)$, et pour mesure de probabilité μ , la mesure de Lebesgue sur $[0, 1]$. Le choix de f n'est peut-être pas optimal, mais semble raisonnable lorsque θ est dans un voisinage de 1. On a :

$$\begin{aligned} F^\theta(x) &= \int_{-\infty}^{-\sqrt{\log(1/x)}} \sqrt{\frac{4\theta}{\pi}} \exp(-\theta z^2) dz, \\ \partial F^\theta(x) &= \frac{1}{2\theta} F^\theta(x) - \int_{-\infty}^{-\sqrt{\log(1/x)}} z^2 \sqrt{\frac{4\theta}{\pi}} \exp(-\theta z^2) dz, \dots, \end{aligned}$$

et l'on vérifie facilement que les conditions (H1)-(H3), (H5), sont satisfaites, montrons que (H4) est aussi vérifiée.

Il faut prouver que $\forall p > 1$, $\mu^\theta(|G_x^\theta|^{2p}) < \infty$ et $\mu^\theta(|\dot{G}_x^\theta|^{2p}) < \infty$, où la fonction G_x^θ est solution de l'équation de Poisson $\mathcal{L}^\theta G_x^\theta = -\tilde{f}_x^\theta$, \mathcal{L}^θ étant le générateur infinitésimal de la diffusion $(X_t^\theta)_{t \geq 0}$.

Dans ce cas particulier, l'on a :

$$\dot{G}_x^\theta(z) = -2 \exp(\theta z^2) \int_{\infty}^z \tilde{f}_x^\theta(u) \exp(-\theta u^2) du = 2 \exp(\theta z^2) \int_z^{\infty} \tilde{f}_x^\theta(u) \exp(-\theta u^2) du.$$

D'autre part les inégalités bien connues sur la fonction de répartition de la loi normale nous permettent d'obtenir l'encadrement suivant, $\forall \theta \in \Theta = (a, b), \forall z \geq 0$:

$$\frac{1}{\theta} \frac{\exp(-\theta z^2)}{z + \sqrt{z^2 + \frac{2}{\theta}}} \leq \int_z^{\infty} \exp(-\theta u^2) du \leq \frac{1}{\theta} \frac{\exp(-\theta z^2)}{z + \sqrt{z^2 + \frac{1}{\theta}}} \leq \frac{1}{a} \frac{\exp(-\theta z^2)}{z + \sqrt{z^2 + \frac{1}{b}}}$$

Par conséquent, comme $\forall u, |\tilde{f}_x^\theta(u)| \leq 1$, on a $|\dot{G}_x^\theta(z)| \leq \frac{C_1}{|z| + C_2}$ pour certaines constantes $C_1, C_2 > 0$ qui dépendent uniquement de a, b , et donc, la première partie de l'hypothèse (H4) est vérifiée. De la même manière, pour $z \geq 0$ $|G_x^\theta(z)| \leq \int_0^z |\dot{G}_x^\theta(u)| du \leq \int_0^z \frac{C_1}{u + C_2} du \leq C_1 \log\left(\frac{|z|}{C_2} + 1\right)$, et donc $\forall p > 1$, $\mu^\theta(|G_x^\theta(z)|^{2p}) < \infty$, la seconde partie de l'hypothèse (H4) est vérifiée.

Nous avons vu que dans ces conditions, l'estimateur de la distance minimale est fortement consistant avec une vitesse de convergence polynomiale de tout ordre, et qu'il admet un développement asymptotique en puissance de $T^{-1/2}$ qui montre en particulier sa normalité asymptotique.

Le processus empirique $(Z_T^{\theta_0}(x), x \in [0, 1])$, dans le cas où l'observation provient du processus $(X_t^{\theta_0})$, converge vers un processus $(Z^{\theta_0}(x), x \in [0, 1])$ de covariance :

$$c^{\theta_0}(x, y) = 4\sqrt{\frac{\theta_0}{\pi}} \int \exp(\theta_0 u^2) \tilde{g}_x^{\theta_0}(u) \tilde{g}_y^{\theta_0}(u) du, \quad x, y \in [0, 1]$$

où l'on a posé $\tilde{g}_x^{\theta_0}(u) = \int_{-\infty}^u \tilde{f}_x^{\theta_0}(v) \exp(-\theta_0 v^2) dv$.

On calcule numériquement la covariance du processus $(Z^{\theta_0}(x), x \in [0, 1])$ sur une grille de points $(x, y) \in [0, 1]^2$. Cette covariance s'avère difficile à évaluer pour certaines valeurs de θ . La variance γ_{θ_0} de la loi limite, lorsque $T \rightarrow \infty$, de $\sqrt{T}(\theta_T^* - \theta_0)$ est donnée par la formule :

$$\gamma_{\theta_0} = \frac{1}{\|\partial F^{\theta_0}(\cdot)\|^4} \int_0^1 \int_0^1 c^{\theta_0}(x, y) \partial F^{\theta_0}(y) \partial F^{\theta_0}(x) dy dx$$

que l'on peut évaluer numériquement.

Nous avons effectué des essais numériques avec les paramètres suivants :

Schéma de discrétisation : Milshtein (ordre 2)

Durée de la simulation : 1000

Pas de discrétisation en temps : $3.83 \cdot 10^{-3}$

Pas de discrétisation en x (covariance) : 0.005

θ : 1.0

On trouve que la variance asymptotique de l'estimateur pour la valeur choisie de θ ($\theta = 1.0$) vaut approximativement 2.25. Pour $\theta = 0.1$, elle vaut approximativement 0.24. Nous avons effectué d'autres essais numériques avec différentes valeurs de θ allant de 0.5 à 10 et nous trouvons un rapport $\frac{\gamma_{\theta}}{\theta}$ à peu près constant et proche de 2.4. A titre indicatif, la variance de l'estimateur du maximum de vraisemblance de θ , qui est asymptotiquement efficace vaut 2θ .

Les figures 1 et 2 ci-après montrent des comportements typiques de cet estimateur.

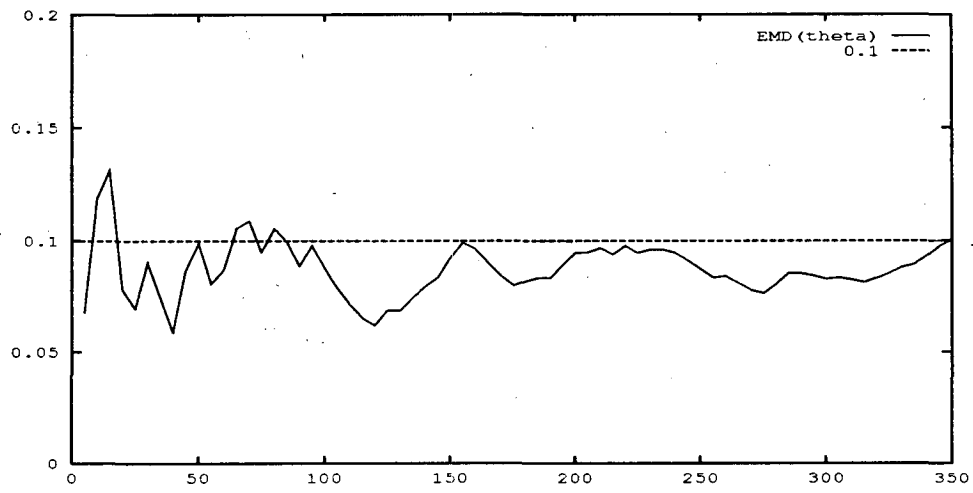


Figure 1: *Estimation de $\theta = 0.1$ sur des données simulées.*

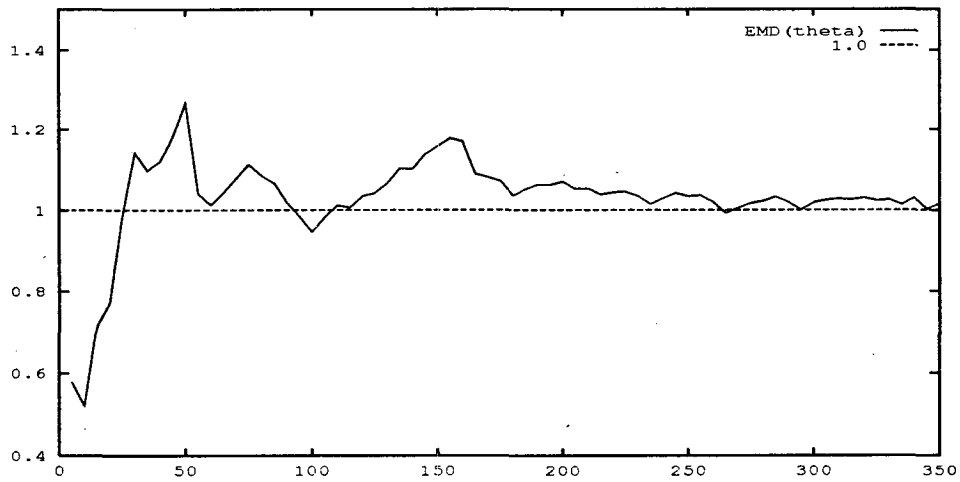


Figure 2: *Estimation de $\theta = 1.0$ sur des données simulées.*

4.2 Exemple 2 : estimation d'un paramètre de diffusion

Nous considérons maintenant un cas où le paramètre θ à estimer est dans le coefficient de diffusion. Plus précisément, soit $\mathcal{F}_\theta = \{(X_t^\theta)_{t \geq 0}, \theta \in \Theta = (a, b), a, b > 0\}$ une famille de processus de diffusion vérifiant les équations différentielles stochastiques

$$dX_t^\theta = -X_t^\theta dt + \theta dW_t$$

On peut facilement montrer que $\forall \theta > 0$, (X_t^θ) est récurrent positif, de mesure invariante

$$d\mu^\theta(x) = \frac{1}{\theta\sqrt{\pi}} \exp\left(-\frac{x^2}{\theta^2}\right) dx.$$

Nous choisissons pour $f : \mathbb{R} \rightarrow [0, 1]$, la fonction $f(x) = \exp(-x^2)$, et pour mesure de probabilité μ , la mesure de Lebesgue sur $[0, 1]$.

On a :

$$\begin{aligned} F^\theta(x) &= \int_{-\infty}^{-\sqrt{\log(1/x)}} \frac{2}{\theta\sqrt{\pi}} \exp\left(-\frac{z^2}{\theta^2}\right) dz, \\ \partial F^\theta(x) &= -\frac{1}{\theta} F^\theta(x) + \frac{2}{\theta^3} \int_{-\infty}^{-\sqrt{\log(1/x)}} z^2 \frac{2}{\theta\sqrt{\pi}} \exp\left(-\frac{z^2}{\theta^2}\right) dz, \dots, \end{aligned}$$

Nous avons effectué des essais numériques avec les paramètres suivants :

Schéma de discrétisation : Milshtein (ordre 2)

Durée de la simulation : 100

Pas de discrétisation en temps : $3.83 \cdot 10^{-3}$

Pas de discrétisation en x (covariance) : 0.005

θ : 1.0

On trouve que la variance asymptotique de l'estimateur θ_T^* pour $\theta = 1.0$ vaut approximativement 0.26. Sur la figure 3 ci-après nous montrons le comportement de cet estimateur.

4.3 Exemple 3 : estimation d'un paramètre de translation

Nous considérons maintenant un processus de Ornstein-Uhlenbeck général, c'est-à-dire la famille $\mathcal{F}_{\bar{\theta}} = \{(X_t^{\bar{\theta}})_{t \geq 0}, \bar{\theta} = (c, \theta, \sigma) \in \Theta = \prod_{i=1..3} (a_i, b_i), a_i, b_i > 0\}$ de processus de diffusion vérifiant les équations différentielles stochastiques

$$dX_t^{\bar{\theta}} = c(\theta - X_t^{\bar{\theta}})dt + \sigma dW_t$$

On montre que $\forall c > 0$, $(X_t^{\bar{\theta}})$ est récurrent positif, de mesure invariante

$$d\mu^{\bar{\theta}}(x) = \sqrt{\frac{c}{\pi\sigma^2}} \exp\left(-\frac{c}{\sigma^2}(x - \theta)^2\right) dx.$$

Nous supposons connu le paramètre de rappel c et le coefficient de diffusion σ , et nous allons estimer le paramètre de translation θ . Nous choisissons pour $f : \mathbb{R} \rightarrow [0, 1]$, la fonction $f(x) = \frac{e^x}{e^x + e^{-x}}$, et pour mesure de probabilité μ , la mesure de Lebesgue sur $[0, 1]$. On a :

$$\begin{aligned} F^{\bar{\theta}}(x) &= \int_{-\infty}^{\log(\sqrt{\frac{x}{1-x}})} \sqrt{\frac{c}{\pi\sigma^2}} \exp\left(-\frac{c}{\sigma^2}(z - \theta)^2\right) dz, \\ \partial_{\theta} F^{\bar{\theta}}(x) &= \int_{-\infty}^{\log(\sqrt{\frac{x}{1-x}})} \frac{2c}{\sigma^2}(z - \theta) \sqrt{\frac{c}{\pi\sigma^2}} \exp\left(-\frac{c}{\sigma^2}(z - \theta)^2\right) dz. \end{aligned}$$

Les essais numériques (figure 4) ont été effectués avec les paramètres suivants :

Schéma de discrétisation : Milshtein (ordre 2)

Durée de la simulation : 1000

Pas de discrétisation en temps : $3.83 \cdot 10^{-3}$

Pas de discrétisation en x (covariance) : 0.005

Valeurs des paramètres : $c = 0.8, \theta = 0.1, \sigma = 1.0$

On trouve numériquement que la variance de l'EMD de θ vaut 1.75. On rappelle que, dans ce cas, l'estimateur du maximum de vraisemblance du paramètre θ est asymptotiquement efficace avec la variance asymptotique $\frac{\sigma^2}{c^2} = 1.56$ pour les valeurs de c, θ et σ choisies. On voit que dans ce cas l'estimateur de la distance minimale du paramètre de translation est proche de l'optimalité.

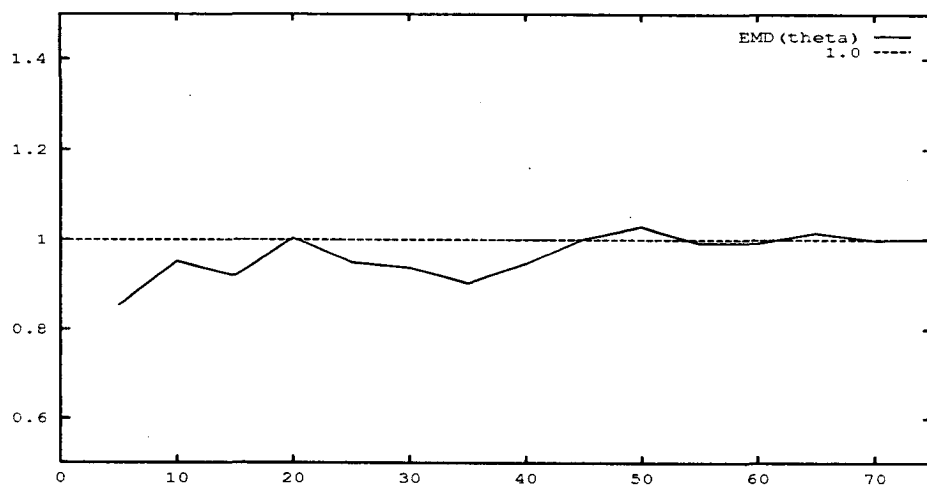


Figure 3: *Estimation de $\theta = 1.0$ sur des données simulées.*

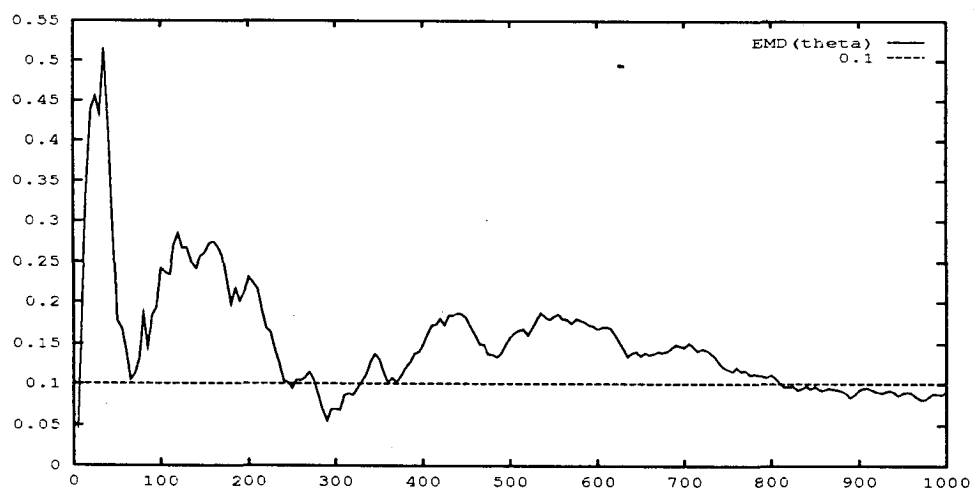


Figure 4: *Estimation de $\theta = 0.1$ sur des données simulées.*

5 Cas des modèles mal spécifiés

Nous allons dans ce paragraphe nous intéresser au comportement de l'estimateur de la distance minimale dans le cas où le modèle est mal spécifié, c'est-à-dire lorsque l'observation est une réalisation d'un processus de diffusion ergodique n'appartenant pas à la famille de diffusions supposée.

Nous considérons à nouveau la famille de processus de diffusion ergodiques $\mathcal{F}_\theta = \{(X_t^\theta)_{t \geq 0}\}$ donnée par les équations différentielles stochastiques (2.1), et nous conservons toutes les hypothèses que nous avons faites sur ce modèle au paragraphe 2.

Supposons maintenant que l'observation est une réalisation du processus de diffusion ergodique $(X_t)_{t \geq 0}$ solution de l'équation différentielle stochastique suivante :

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad t \geq 0 \quad (5.10)$$

où $b(x)$ est une fonction borélienne, bornée sur tout compact de \mathbb{R} , $\sigma^2(x)$ est une fonction continue, strictement positive. Soit $Q_{b,\sigma}^\nu$ la loi induite par (X_t) lorsque sa loi initiale est ν sur $(C(\mathbb{R}^+; \mathbb{R}), \mathcal{B})$, et soit $\mu^{b,\sigma}$ la loi invariante associée.

Nous allons poser quelques définitions et hypothèses dont nous aurons besoin par la suite. On note \mathcal{L} l'opérateur infinitésimal de la diffusion (X_t) .

$$\mathcal{L} = b(x) \frac{\partial}{\partial x} + \frac{\sigma^2(x)}{2} \frac{\partial^2}{\partial x^2}$$

Posons $s(x) = \exp(-\int_0^x \frac{2b(z)}{\sigma^2(z)} dz)$, et $S(x) = \int_0^x s(z) dz$. Nous supposons que la diffusion est récurrente positive, en posant comme d'habitude $\lim_{x \rightarrow \pm\infty} S(x) = \pm\infty$ et $\int \frac{1}{\sigma^2(x)s(x)} dx = M^{-1} < \infty$. La densité de la loi invariante $\mu^{b,\sigma}$ s'écrit alors $\frac{M}{\sigma^2(x)s(x)}$. Soit G la fonction de répartition associée à la mesure invariante $\mu^{b,\sigma}$. Posons $\tilde{f}_x = \chi_{[0,x]} \circ f - \int \chi_{[0,x]} \circ f d\mu^{b,\sigma}$, et considérons le processus différence empirique $(\xi_T(x), 0 \leq x \leq 1)$ et le processus empirique normalisé $(Z_T(x), 0 \leq x \leq 1)$ associé à (X_t) :

$$\begin{aligned} \xi_T(x) &= F_T(x) - G(x) \\ Z_T(x) &= \sqrt{T}(F_T(x) - G(x)) \\ &= \frac{1}{\sqrt{T}} \int_0^T \tilde{f}_x(\pi_t) dt. \end{aligned}$$

Définissons θ^* comme suit :

$$\theta^* \in \arg \min_{\theta \in \Theta} \|G(\cdot) - F^\theta(\cdot)\|$$

Définissons d'autre part

$$h(\eta) = \min_{|\theta - \theta^*| > \eta} \|G(\cdot) - F^\theta(\cdot)\| - \|G(\cdot) - F^{\theta^*}(\cdot)\|,$$

et faisons l'hypothèse d'identifiabilité :

(H6) $\forall \eta > 0, h(\eta) > 0$.

5.1 Convergence

Nous faisons l'hypothèse suivante :

(H7) $\forall x \in [0, 1]$, l'équation de Poisson $\mathcal{L}u = -\tilde{f}_x$ admet une solution $u = K_x$ vérifiant $\forall p$ entier pair, positif, $\mu^{b,\sigma}(\sigma^p \dot{K}_x(\cdot)^p) < \infty, \mu^{b,\sigma}(K_x^p) < \infty$.

On montre alors le

Théorème 5.10 *Supposons vérifiées les hypothèses (H6), (H7), alors $\forall p$ entier pair, positif, $\exists C_p > 0, \forall \eta > 0$:*

$$Q_{b,\sigma}^{\mu^{b,\sigma}}(|\theta_T^* - \theta^*| > \eta) \leq \frac{C_p}{T^{p/2} h(\eta)^p}, \quad \forall T > 0.$$

Preuve Sous $Q_{b,\sigma}^{\mu^{b,\sigma}}$, nous avons les inégalités suivantes :

$$\begin{aligned} \|\xi_T(\cdot)\| + \|G(\cdot) - F^{\theta^*}(\cdot)\| &\geq \min_{|\theta - \theta^*| \leq \eta} \|\xi_T^\theta(\cdot)\|, \\ \min_{|\theta - \theta^*| > \eta} \|\xi_T^\theta(\cdot)\| &\geq \min_{|\theta - \theta^*| > \eta} \|G(\cdot) - F^\theta(\cdot)\| - \|\xi_T(\cdot)\| \end{aligned}$$

Par conséquent,

$$\left\{ \min_{|\theta - \theta^*| \leq \eta} \|\xi_T^\theta(\cdot)\| > \min_{|\theta - \theta^*| > \eta} \|\xi_T^\theta(\cdot)\| \right\} \subset \left\{ \|\xi_T(\cdot)\| > \frac{h(\eta)}{2} \right\}.$$

c'est-à-dire,

$$\{|\theta_T^* - \theta^*| > \eta\} \subset \left\{ \|\xi_T(\cdot)\| > \frac{h(\eta)}{2} \right\}.$$

La démonstration suit ensuite pas à pas celle du théorème 2.2 en utilisant l'inégalité de Markov et la représentation du processus $Z_T(x)$

$$Z_T(x) = \frac{1}{\sqrt{T}} \int_0^T \tilde{f}_x(\pi_t) dt = \frac{1}{\sqrt{T}} \left\{ K_x(\pi_0) - K_x(\pi_T) + \int_0^T (\sigma \dot{K}_x)(\pi_t) dW_t \right\}$$

□

5.2 Normalité asymptotique

Dans ce paragraphe, nous montrons la normalité asymptotique de l'estimateur de la distance minimale pour le cas mal spécifié. Nous faisons les hypothèses suivantes, dont la première assure la convergence en loi du processus empirique Z_T associé à l'observation (5.10) vers un processus gaussien continu, centré Z dont nous notons $c(x, y)$ la covariance (voir Fournié [5], th. II.2.13) :

(H8) $\forall x \in [0, 1]$, l'équation de Poisson $\mathcal{L}u = -\tilde{f}_x$ admet une solution $u = K_x$ vérifiant $\mu^{b,\sigma}(\sigma^2 K_x(\cdot)^2) < \infty$.

(H9) $J_{\theta^*} = \left(\partial F^{\theta^*}(\cdot), {}^t \partial F^{\theta^*}(\cdot) \right) + \left(F^{\theta^*}(\cdot) - G(\cdot), \partial^2 F^{\theta^*}(\cdot) \right)$ est inversible.

Théorème 5.11 *Supposons vérifiées les hypothèses (H2), (H3), (H6), (H8) et (H9) alors, sous la loi $Q'_{b,\sigma}$:*

$$\sqrt{T}(\theta_T^* - \theta^*) \implies \mathcal{N}(0, \Gamma_{\theta^*}), \quad T \rightarrow \infty$$

où $\Gamma_{\theta^*} = J_{\theta^*}^{-1} M_{\theta^*} J_{\theta^*}^{-1}$ et $M_{\theta^*} = (M_{\theta^*}^{ij})$, avec

$$M_{\theta^*}^{ij} = \int_0^1 \int_0^1 c(x, y) \partial_i F^{\theta^*}(y) \partial_j F^{\theta^*}(x) d\mu(y) d\mu(x).$$

et $J_{\theta^*} = \left(\partial F^{\theta^*}(\cdot), {}^t \partial F^{\theta^*}(\cdot) \right) + \left(F^{\theta^*}(\cdot) - G(\cdot), \partial^2 F^{\theta^*}(\cdot) \right)$

Preuve Nous avons par définition

$$\min_{\theta \in \Theta} \|F_T(\cdot) - F^\theta(\cdot)\| = \|F_T(\cdot) - F^{\theta_T^*}(\cdot)\|$$

et par conséquent l'hypothèse de différentiabilité (H2) montre que les points θ_T^* vérifient le système de k équations :

$$\begin{aligned} & \left(F_T(\cdot) - F^{\theta_T^*}(\cdot), \partial_i F^{\theta_T^*}(\cdot) \right) \\ &= \left((F_T(\cdot) - G(\cdot)) - (F^{\theta_T^*}(\cdot) - G(\cdot)), \partial_i F^{\theta_T^*}(\cdot) \right) = 0, \quad i = 1 \dots k \end{aligned}$$

Par conséquent, en appliquant le théorème des accroissements finis aux fonctions $f_i(\theta) = \left(F^\theta(\cdot) - G(\cdot), \partial_i F^\theta(\cdot) \right)$ au point θ^* , et en remarquant que $f_i(\theta^*) = 0$, pour $i = 1 \dots k$, nous obtenons le système suivant :

$$\begin{aligned} \left(F_T(\cdot) - G(\cdot), \partial_i F^{\theta_T^*}(\cdot) \right) &= \sum_{j=1}^k (\theta_{jT}^* - \theta_j^*) \left[\left(\partial_j F^{\theta_{iT}^*}(\cdot), \partial_i F^{\theta_{iT}^*}(\cdot) \right) \right. \\ &\quad \left. + \left(F^{\theta_{iT}^*}(\cdot) - G(\cdot), \partial_{ij}^2 F^{\theta_{iT}^*}(\cdot) \right) \right] \\ &= \sum_{j=1}^k (\theta_{jT}^* - \theta_j^*) J_{\theta_{iT}^*}^{ij}, \quad i = 1 \dots k \end{aligned}$$

où $\tilde{\theta}_{iT}$ appartient au segment de droite $[\theta^*, \theta_T^*]$, pour $i = 1 \dots k$. Comme θ_T^* converge en probabilité vers θ^* , et que J_{θ^*} est inversible par l'hypothèse (H9), pour tout T assez grand on peut inverser la matrice $J_{\tilde{\theta}_{iT}} = (J_{\tilde{\theta}_{iT}}^{ij})_{i,j}$ (cf. une argumentation similaire dans la preuve du théorème 3.4 dans [6]) et en multipliant chaque membre par \sqrt{T} , nous obtenons:

$$\sqrt{T}(\theta_T^* - \theta^*) = J_{\tilde{\theta}_{iT}}^{-1} (Z_T(\cdot), \partial F^{\theta_T^*}(\cdot))$$

Le théorème 2.3 et l'hypothèse (H8) entraîne alors la convergence en loi, sous $Q_{b,\sigma}^\nu$, lorsque $T \rightarrow \infty$, de $\sqrt{T}(\theta_T^* - \theta^*)$ vers la variable gaussienne centrée $J_{\theta^*}^{-1} (Z(\cdot), \partial F^{\theta^*}(\cdot))$ dont la matrice de covariance est donnée dans l'énoncé du théorème. \square

Remarque 1 : on notera que lorsque le modèle est bien spécifié, c'est-à-dire lorsque $\|G(\cdot) - F^{\theta^*}(\cdot)\| = 0$, on retrouve le théorème de normalité asymptotique 2.4.

Remarque 2 : l'estimateur de θ , dans le cas mal spécifié peut converger plus vite vers la valeur θ^* minimisant $\|G(\cdot) - F^\theta(\cdot)\|$, que dans le cas bien spécifié vers la vraie valeur.

5.3 Exemple 4

Nous reprenons l'exemple de la famille $\mathcal{F}_{\bar{\theta}} = \{(X_t^{\bar{\theta}})_{t \geq 0}, \bar{\theta} = (c, \theta, \sigma) \in \Theta = (a, b)^3, a, b > 0\}$ de processus de Ornstein-Uhlenbeck ergodiques vérifiant les équations différentielles stochastiques

$$dX_t^{\bar{\theta}} = c(\theta - X_t^{\bar{\theta}})dt + \sigma dW_t \quad (5.11)$$

L'observation n'appartient pas à cette famille, mais vérifie l'équation différentielle stochastique :

$$dX_t = -X_t^3 dt + dW_t \quad (5.12)$$

On montre facilement que les hypothèses du théorème 5.10 sont vérifiées.

La fonction $f : \mathbb{R} \rightarrow [0, 1]$ est la même que dans l'exemple 3, $f(x) = \frac{e^x}{e^x + e^{-x}}$.

Nous supposons que le coefficient de diffusion σ est connu, et nous identifions simultanément les paramètres c et θ pour ce cas mal spécifié.

Les essais numériques qui suivent ont été effectuées avec les paramètres suivants :

Schéma de discrétisation : Milshtein (ordre 2)

Durée de la simulation : 1000

Pas de discrétisation en temps : $3.83 \cdot 10^{-3}$

Pas de discrétisation en x (covariance) : 0.005

Valeurs des paramètres : $\sigma = 1.0$

On voit sur les graphes 5 et 6 qui suivent que les estimateurs de c et θ convergent vers les valeurs qui minimisent la distance L_{μ}^2 entre la famille des lois invariantes du modèle (5.11) et la loi invariante associée du processus solution de (5.12), qui valent dans notre cas $c^* \simeq 0.94$ et $\theta^* \simeq 0$.

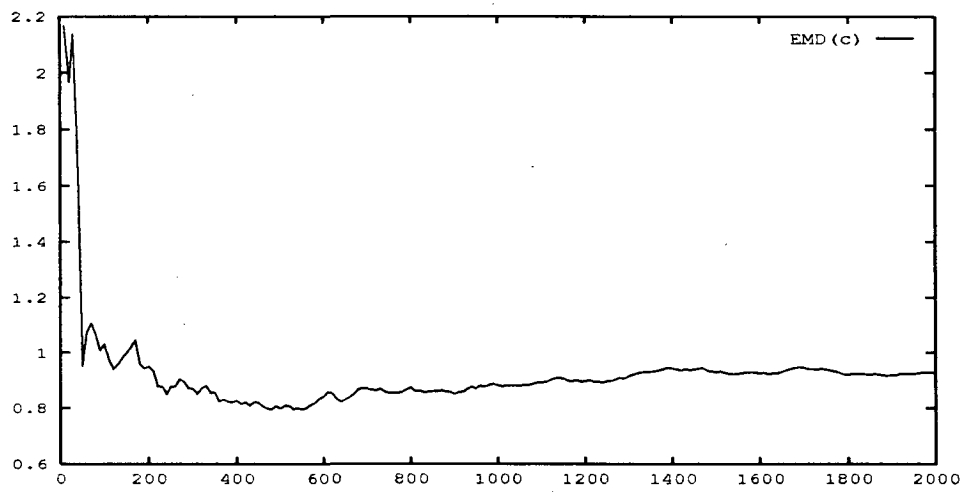


Figure 5: *Estimation de c sur des données simulées.*

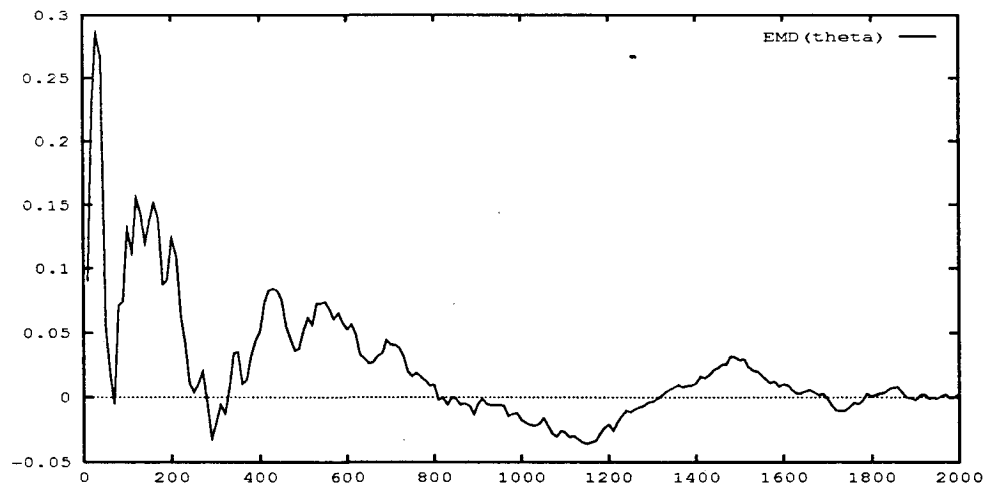


Figure 6: *Estimation de θ sur des données simulées.*

5.4 Identification d'un modèle de taux d'intérêt

Nous nous intéressons maintenant à l'identification d'un modèle paramétrique d'évolution des taux d'intérêt proposé par J.C. Cox, J.E. Ingersoll et S.A. Ross dans [2] (modèle C.I.R.) défini par les équations différentielles stochastiques :

$$dX_t = c(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t, \quad X_0 > 0 \quad (5.13)$$

où c , θ et σ sont des paramètres réels représentant respectivement le coefficient de rappel (difficile à interpréter économiquement), la moyenne et la volatilité du taux d'intérêt.

Nous avons mené dans [6] une étude détaillée des estimateurs du maximum de vraisemblance des coefficients c et θ , qui sont des estimateurs asymptotiquement efficaces pour ce modèle. Mais nous avons vu d'autre part que l'estimation non paramétrique de la dérive et de la fonction de diffusion ne permettait pas d'affirmer ou d'infirmer l'adéquation du modèle C.I.R. à des données réelles de taux d'intérêt. Les tests de modèles proposés dans [5] n'apportaient pas non plus de réponse définitive à cette question. Nous sommes donc clairement dans le cadre de l'estimation paramétrique pour des modèles mal spécifiés, et il semble judicieux d'utiliser dans ce cas des méthodes d'estimation robustes plutôt que des méthodes efficaces lorsque le modèle est parfaitement connu ¹.

Nous avons vu au paragraphe précédent que l'estimateur du minimum de distance présente une certaine propriété de robustesse, différente de celle de l'EMV, et converge sous des hypothèses plus larges. Il couvre donc une classe de modèles plus vaste, et il semblait intéressant d'étudier son comportement sur des données réelles. Nous l'avons mis en oeuvre sur 11 ans de relevés journaliers sur le marché monétaire français, de Janvier 1978 à Janvier 1989, et nous avons obtenu les résultats suivants.

Le paramètre de volatilité σ a été estimé par la méthode de la variation quadratique (voir chapitre I paragraphe 2.1). Il est supposé connu et égal à 0.1.

Nous avons choisi pour fonction $f : \mathbb{R} \rightarrow [0, 1]$, $f(x) = \exp(-x)$. Les valeurs des taux d'intérêt qui nous intéressent varient entre 0.05 et 0.18, et l'on peut considérer que f est linéaire dans cet intervalle. on ne privilégie donc pas certaines valeurs plus que d'autres lorsque nous mettons en oeuvre notre algorithme de minimisation.

¹Ian W. MacKeague [14] a montré une certaine propriété de robustesse de l'estimateur du maximum de vraisemblance pour un paramètre unidimensionnel, aisément généralisable au cas multidimensionnel. Plus précisément, si l'observation est une réalisation du processus $(X_t)_{t \geq 0}$, stationnaire, de loi invariante μ , vérifiant l'E.D.S :

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 \simeq \mu \quad (5.14)$$

Notons $Q_b^{\mu, \sigma}$ la loi induite sur $(C(\mathbb{R}^+; \mathbb{R}), \mathcal{B})$. Soit $(\hat{c}_T, \hat{\theta}_T)$ l'estimateur du maximum de vraisemblance de (c, θ) dans le modèle (5.13). Sous des hypothèses assez générales, $(\hat{c}_T, \hat{\theta}_T) \rightarrow (c^*, \theta^*)$, $Q_b^{\mu, \sigma}$ -p.s, lorsque $T \rightarrow \infty$, où (c^*, θ^*) est l'unique minimum en (c, θ) de la fonction $E \left[\frac{b(X) - c(X - \theta)}{\sigma\sqrt{X}} \right]^2$, où X est une variable aléatoire de loi μ . Ces hypothèses ne sont pas toujours vérifiées.

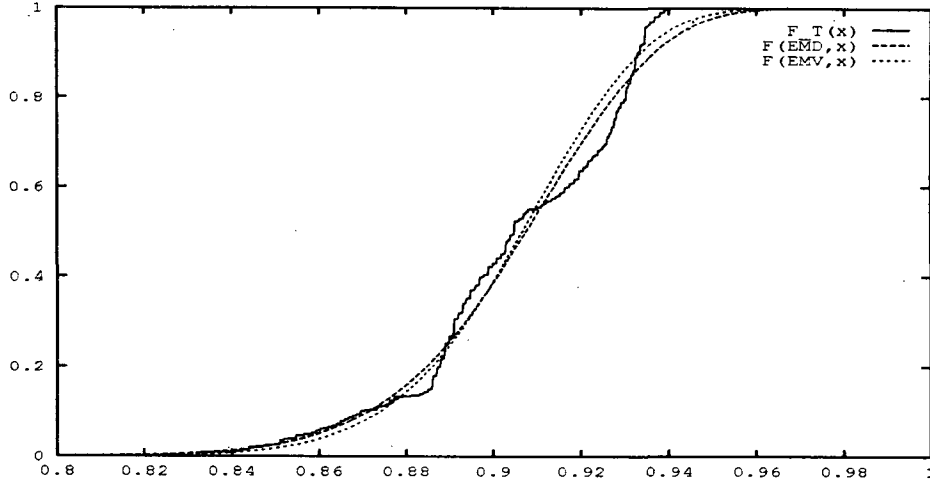


Figure 7: $F_T, F(c_T^*, \theta_T^*), F(\hat{c}_T, \hat{\theta}_T)$ sur données réelles.

Posons $\delta(c, \theta) = \|F_T(\cdot) - F^{c, \theta}(\cdot)\|_{L_\mu^2}$, où μ est la mesure de Lebesgue. Nous avons obtenu les résultats suivants :

- EMD : $(c_T^*, \theta_T^*) = (0.616, 0.098)$ avec $\delta(0.616, 0.098) = 1.7 \cdot 10^{-4}$.
- EMV : $(\hat{c}_T, \hat{\theta}_T) = (0.715, 0.102)$ avec $\delta(0.715, 0.102) = 2.1 \cdot 10^{-4}$.

On voit qu'en estimant par la méthode du maximum de vraisemblance, la fonction de répartition associée aux estimations pour le modèle (5.13) n'est pas très loin de la fonction de répartition empirique.

La figure 7 montre les graphes des fonctions $F_T, F(c_T^*, \theta_T^*)$ et $F(\hat{c}_T, \hat{\theta}_T)$.

6 Théorème de limite centrale fonctionnelle pour le processus empirique ajusté

Nous faisons un petit retour sur le test de Kolmogorov-Smirnov avec paramètres estimés proposé dans [5], chapitre II. Plus précisément, nous allons donner la loi asymptotique, sous Q_θ' , du processus $Z_T^{\theta_T^*}(x) = \sqrt{T} (F_T(x) - F^{\theta_T^*}(x))$, $\forall x \in [0, 1]$, où θ_T^* est l'estimateur de la distance minimale. L'avantage de cette procédure réside dans le fait que l'estimateur de la distance minimale permet d'identifier des paramètres figurant indifféremment dans la dérive ou dans le coefficient de diffusion de l'observation. En considérant le suprémum du processus gaussien limite de $Z_T^{\theta_T^*}$ ou bien une statistique de type Chi deux (voir [5], chapitre II), nous obtenons un véritable test de modèle de processus de diffusion ergodiques réels dans une asymptotique grand temps (dans le cas particulier, bien sûr, où la loi invariante permet l'identification totale du modèle).

Théorème 6.12 Supposons vérifiées les hypothèses (H1)-(H3) et les hypothèses du théorème 2.3, alors $\forall \theta_0 \in \Theta$, on a sous la loi $Q_{\theta_0}^\nu$:

$$Z_T^{\theta_0^*} \Longrightarrow Z_{\theta_0}^*$$

dans $D[0,1]$, lorsque $T \rightarrow \infty$, où $Z_{\theta_0}^*$ est un processus gaussien à trajectoires $Q_{\theta_0}^{\nu,\sigma}$ -p.s. continues de moyenne nulle et de covariance

$$\begin{aligned} c_{\theta_0}^*(x, y) &= c^{\theta_0}(x, y) - {}^t \partial F^{\theta_0}(x) J_{\theta_0}^{-1} \int_0^1 c^{\theta_0}(z, y) \partial F^{\theta_0}(z) d\mu(z) \\ &\quad - {}^t \partial F^{\theta_0}(y) J_{\theta_0}^{-1} \int_0^1 c^{\theta_0}(x, z) \partial F^{\theta_0}(z) d\mu(z) + {}^t \partial F^{\theta_0}(x) \Gamma_{\theta_0} \partial F^{\theta_0}(y). \end{aligned}$$

avec les notations du théorème 2.4.

Preuve La preuve est quasiment la même que celle du théorème II.3.23 dans [5]. En effet, on a :

$$Z_T^{\theta_0^*}(x) = Z_T^{\theta_0}(x) - \sqrt{T} (F_T^{\theta_0^*}(x) - F^{\theta_0}(x)).$$

Nous savons que (voir le paragraphe 2 de ce chapitre) :

$$\begin{aligned} \sqrt{T}(\theta_T^* - \theta_0) &= J_{\theta_0}^{-1} (\partial F^{\theta_0}(\cdot), Z_T^{\theta_0}(\cdot)) + \varepsilon(T) \\ &= h_{\theta_0}(Z_T^{\theta_0}(\cdot)) + \varepsilon(T) \end{aligned}$$

où $h : D[0,1] \rightarrow \mathbb{R}^k$ est continue, et $\varepsilon(T) \rightarrow 0, T \rightarrow \infty$.

Considérons la famille de fonctions $(g_t^{\theta_0})_{t>0}$ et la fonction g^{θ_0} définies comme suit :

$$\begin{aligned} g_t^{\theta_0} &: D[0,1] \longrightarrow D[0,1] \\ g_t^{\theta_0}(X(\cdot)) &= X(\cdot) - \sqrt{t} \left(F^{\theta_0+t^{-1/2}(h_{\theta_0}(X(\cdot))+\varepsilon(t))}(\cdot) - F^{\theta_0}(\cdot) \right) \\ g^{\theta_0} &: D[0,1] \longrightarrow D[0,1] \\ g^{\theta_0}(X(\cdot)) &= X(\cdot) - \langle \partial F^{\theta_0}(\cdot), h_{\theta_0}(X(\cdot)) \rangle \end{aligned}$$

Le théorème 5.5 dans Billingsley [1] et les hypothèses du théorème montre que sous $Q_{\theta_0}^\nu$, lorsque $T \rightarrow \infty$, $g_T^{\theta_0}(Z_T^{\theta_0}(\cdot)) \Longrightarrow g^{\theta_0}(Z^{\theta_0}(\cdot))$, c'est-à-dire $Z_T^{\theta_0^*}(\cdot) \Longrightarrow Z_{\theta_0}^*(\cdot)$.

Les mêmes arguments que ceux utilisés dans la preuve du théorème II.3.23 dans [5], entraînent que $Z_{\theta_0}^*(\cdot)$ est un processus gaussien centré à trajectoires $Q_{\theta_0}^\nu$ -p.s. continues, dont nous identifions facilement la covariance $c_{\theta_0}^*(x, y)$. \square

7 Conclusion

Nous avons étudié un estimateur d'un paramètre pour des processus de diffusion ergodiques obtenu par la méthode de la distance minimale. Cet estimateur est basé sur la distance minimale L_μ^2 entre la fonction de répartition empirique et les fonctions de répartition théoriques données par le modèle, renvoyées sur l'intervalle $[0, 1]$ par l'intermédiaire d'une fonction $f : \mathbb{R} \rightarrow [0, 1]$ (dont le choix est important dans les applications). Nous avons montré sa consistance forte et sa normalité asymptotique, et dans le cas de la dimension un, nous avons donné son développement complet sur un événement de probabilité croissante vers 1 avec le temps d'observation. Nous avons d'autre part évalué ses performances de manière approfondie sur des modèles simples de diffusion (processus d'Ornstein-Uhlenbeck), pour différents types de paramètres : paramètres de translation et de rappel dans la dérive, paramètre de volatilité dans le terme de diffusion. Nous avons ensuite étudié le cas d'un modèle mal spécifié pour lequel nous avons montré la convergence de l'estimateur de la distance minimale vers la valeur θ^* du paramètre minimisant la distance L_μ^2 entre la famille des fonctions de répartition invariantes du modèle et celle associée à l'observation. Nous avons aussi montré la normalité asymptotique de l'estimateur autour de cette valeur. Un terme correcteur apparaît dans la covariance asymptotique par rapport au cas bien spécifié.

Comme l'on pouvait s'en douter a priori, cet estimateur est en général moins bon que l'estimateur du maximum de vraisemblance, si celui-ci est disponible, lorsque l'on veut identifier un paramètre appartenant au terme de dérive (on montre en effet, sous des hypothèses assez faibles, i.e. la normalité asymptotique locale du rapport de vraisemblance plus quelques autres conditions, voir Kutoyants [8], que l'estimateur du maximum de vraisemblance est asymptotiquement efficace). D'autre part, même dans les cas simples, par exemple le cas linéaire, on ne peut pas obtenir de formule explicite de l'estimateur du minimum de distance. Ceci n'est pas un grand problème car on peut obtenir de bonnes approximations numériques dans la plupart des cas. Par contre, le calcul de la variance de cet estimateur est complexe et peut poser des problèmes numériques.

Cet estimateur possède néanmoins une propriété remarquable d'universalité dans le sens où il permet d'identifier indifféremment des paramètres figurant dans les termes de dérive ou de diffusion, ce que ne permet pas de faire l'estimateur du maximum de vraisemblance.

D'autre part son calcul ne dépend pas fortement du calcul de l'intégrale d'Itô comme l'estimateur du maximum de vraisemblance ou l'estimateur bayésien.

Parfois, son calcul peut être plus facile que celui de l'estimateur du maximum de vraisemblance ou de l'estimateur bayésien.

On notera aussi que dans le cas des modèles mal spécifiés l'estimateur du minimum de distance est convergent pour une classe d'observation plus large que l'estimateur du maximum de vraisemblance, et peut parfois être meilleur que l'estimateur du maximum de vraisemblance. Enfin, on peut utiliser cet estimateur dans les cas où le rapport de

vraisemblance est inconnu, mais où la distribution invariante est connue.

Nous avons terminé cette étude par un retour sur le test de modèle proposé dans [5], chapitre II. Nous donnons un théorème de limite centrale fonctionnelle pour le processus empirique "ajusté" $Z_T^{\theta_T^*}$. En considérant le suprémum du processus gaussien limite de $Z_T^{\theta_T^*}$ ou bien une statistique de type Chi deux (voir [5], chapitre II), nous pouvons obtenir un test de modèle de processus de diffusion ergodiques réels dans une asymptotique grand temps. L'avantage de la procédure proposée ici étant que les paramètres à estimer peuvent figurer indifféremment dans la dérive ou bien dans le coefficient de diffusion du modèle.

Références

- [1] P. BILLINGSLEY : *"Convergence of probability measures"*, John Wiley and Sons , 1968.
- [2] J.C. COX, J.E. INGERSOLL & S.A. ROSS : A theory of the term structure of the interest rates , *Econometrica* , vol. 53 , 1985.
- [3] D. DACUNHA-CASTELLE & M. DUFLO : *"Probabilité et statistiques, tome 2"*, Masson , 1983.
- [4] HANS M. DIETZ, Y.A. KUTOYANTS : Asymptotic normality of minimum distance estimator of a parameter of ergodic process , 1992 , preprint.
- [5] E. FOURNIE : Statistiques des diffusions ergodiques avec applications en finance , Thèse de Doctorat de Mathématiques Appliquées de l'Université de Nice , Mai 1993.
- [6] E. FOURNIE & D. TALAY : Application de la statistique des diffusions à un modèle de taux d'intérêt , *Finance* , Vol.12 , 2/1991.
- [7] I. KARATZAS & S.E. SHREVE : *"Brownian motion and stochastic calculus"* , Springer-Verlag , 1988.
- [8] Yu.A. KUTOYANTS : *"Parameter estimation for stochastic processes"* , Heldermann Verlag Berlin , 1984.
- [9] Yu.A. KUTOYANTS : Minimum distance parameter estimation for diffusion type observations , *Note du CRAS* , t.312 , Série I , p. 637-642 , 1991.
- [10] Yu.A. KUTOYANTS : On an identification problem of dynamical system with a small noise , *Soviet. J. of Contemp. Math. Anal.* , 23 , 3 , p. 79-95 , 1988.
- [11] Yu.A. KUTOYANTS : Expansion of a maximum likelihood estimate by diffusion powers , *Theo. of Proba. and Its Appl.* , Vol. XXIX , 3 , p. 465-477.

- [12] Yu.A. KUTOYANTS & P. PILIBOSSIAN : On the minimum uniform metric estimate of parameters of diffusion type processes , Rapp. tech. 142 , L.S.T.A. , Juin 1991.
- [13] Yu.A. KUTOYANTS & P. PILIBOSSIAN : On the minimum L_1 -norm estimate of the parameter of Ornstein-Uhlenbeck process , L.S.T.A. , Avril 1992.
- [14] Ian W. MacKEAGUE : Estimation for Diffusion Processes under Misspecified Models , J. Appl. Prob.21 , p 511-520 (1984).
- [15] P.W. MILLAR : The minimax principle in asymptotic statistical theory , Lectures Notes in Mathematics 976 , Springer-Verlag , Berlin , 1983 , p. 75-265.
- [16] P.W. MILLAR : A general approach to the optimality of minimum distance estimators , Trans.Am.Math.Soc. vol. 286 , num.1 , Nov.1984.
- [17] D. TALAY : Simulation and numerical analysis of S.D.E's : a review , in "*Effective Stochastic Analysis*" , P.Kree & W.Wedig (Eds.), Springer-Verlag , à paraître.
- [18] N. YOSHIDA : Asymptotic expansions of for small diffusions via the theory of Malliavin-Watanabe , Proba. Theo. and Relat. Fields , Vol. 92 , p. 275-311 , 1992.



Unité de Recherche INRIA Sophia Antipolis
2004, route des Lucioles - B.P. 93 - 06902 SOPHIA ANTIPOLIS Cedex (France)

Unité de Recherche INRIA Lorraine Technopôle de Nancy-Brabois - Campus Scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 VILLERS LES NANCY Cedex (France)

Unité de Recherche INRIA Rennes IRISA, Campus Universitaire de Beaulieu 35042 RENNES Cedex (France)

Unité de Recherche INRIA Rhône-Alpes 46, avenue Félix Viallet - 38031 GRENOBLE Cedex (France)

Unité de Recherche INRIA Rocquencourt Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 LE CHESNAY Cedex (France)

EDITEUR

INRIA - Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 LE CHESNAY Cedex (France)

ISSN 0249 - 6399

